Gravitational Waves As Tools For Astrophysics And Cosmology

A DISSERTATION
SUBMITTED TO THE FACULTY OF THE GRADUATE SCHOOL
OF THE UNIVERSITY OF MINNESOTA
BY

Sharan Banagiri

IN PARTIAL FULFILLMENT OF THE REQUIREMENTS
FOR THE DEGREE OF
DOCTOR OF PHILOSOPHY

Professor Vuk Mandic

May, 2021

Acknowledgements

First of all, my heartfelt thanks to my advisor Vuk Mandic. Vuk, my Ph.D. has only been possible with your support, encouragement, and guidance. Thank you for supporting me in exploring my own interests while analyzing my ideas with a critical eye at the same time. You are a role model for a great scientist and a fantastic mentor! I must also thank Michael Coughlin for guiding me with my first gravitational wave projects and for always being willing to answer the many questions I had. Your support has been invaluable.

I also thank all of the Minnesota gravitational-wave group that I had the pleasure of working with all these years. Thanks especially to Andrew Matas, Rich Ormiston, and Pat Meyers for many interesting conversations and discussions. I also thank Andrew for teaching me how to think critically about science, making me a better scientist. I also owe thanks to Rich for reading chapters of my thesis and helping me make it better.

I'm also grateful to Claudia Scarlata for all the support and encouragement with the large-scale structure project and beyond. Thank you also for asking me to be a TA for the astrostatistics course. I enjoyed being a TA but I also learned so much from it.

I was also fortunate to work with several other excellent scientists during my Ph.D., including Marie-Anne Bizouard, Alexander Criswell, James Clark, David Keitel, Paul Lasky, Joe Romano, Lilli Sun, Eric Thrane, Kate Yang, and others. Thank you for everything you have taught me over the years!

My thanks also to everyone at the LIGO Hanford Observatory for making my stay as a fellow there memorable. Robert Schofield and Rick Savage, in particular, have taught me much about the detectors. I also had the privilege of meeting and working with the other wonderful fellows at LHO, especially Dripta Bhattacharjee, Pep Covas, Sundae Chen, Laurence Datrier, Lilli Sun, and Kara Merfeld.

I also thank Prof. Narendra Sahu at IIT Hyderabad for believing in me and helping me transition from engineering to physics at the end of my undergraduate degree.

Thank also to my friends from IIT Hyderabad, Aniket Deshmukh, Kalyan Tej, Ravi Chandra, Khalid Mohammad, Aditya Srinivas, Sachin Rathod, and Siddartha Kamepalli, for several memorable road trips during the past few years and fantastic conservations during them. Thank you to Siddarth Karuka, Gautham Gampa, Harris Ahmed, Swapnil Deshpande, Afroz Mohammad, and Amandeep Gautham for your friendship over the years in Minnesota. Thanks to Harison Wiesman, Xuzhe Ying, Michael Sammon, Nick Mast, and others for the Friday coffee hour conservations in the before times and for the online board games during the pandemic.

None of this would have been possible without the love and support of my family, especially my amma (mom), nanna (dad), and my little brother Saketh. I also owe thanks to my nanna for introducing me to science and astronomy when I was very young and nurturing my interest as I grew up. I would not have been on this path without you!

Finally, a part of this work was done during the COVID-19 pandemic, and so I thank all the health care professionals and essential workers who have been at the forefront of fighting against it.

Dedication

To my family

Abstract

The discovery of gravitational waves by LIGO and Virgo have unveiled a sector of the Universe previously hidden from us. Gravitational waves allow us to detect and observe the dynamics of phenomena usually hidden from electromagnetism probes like binary black hole mergers, providing new tools to study astrophysics and cosmology in the process. As the number of detections increase, the statistics of the mergers are starting to inform us about their progenitor distributions.

This dissertation consists of three parts. First, analyses to detect potential gravitational waves from a post-merger remnant of the binary neutron star merger GW170817 are described, with particular emphasis on the STAMP pipeline-based search targeting gravitational waves from a long-lived remnant. Bayesian parameter estimation techniques for the poorly modeled post-merger signals are then described. A novel likelihood formalism is developed to account for the inaccuracies in models, focusing in particular on the phase evolution of the waveform.

In the second part, techniques are developed, using hierarchical Bayesian modeling to measure N-point correlations of the distributions of black hole mergers, with a focus on two-point correlations. These methods allow us to use black hole mergers as a tool to measure the angular distribution and the large-scale structure of the matter in the Universe. The two-point correlation method is validated with simulations for the angular structure of the mergers in the Universe.

Finally, Bayesian methods are devised to probe anisotropies in the angular distribution of the stochastic gravitational-wave backgrounds and foregrounds in the LISA band. A novel decomposition using Clebsch–Gordan coefficients in the spherical harmonic basis is developed that allows us to infer the anisotropy of arbitrary distributions of gravitational-wave power. This method is employed and tested using different kinds of simulations, including that of the galactic gravitational-wave foreground from galactic white dwarf binaries.

Contents

List of Tables

List of Figures

The story so far:
In the beginning the Universe was created. This has made a lot of people very angry and been widely regarded as a bad move.

Douglas Adams, The Restaurant at the End of the Universe

Chapter 1

Gravitational waves

1.1 General Relativity

General relativity is one of the foundational pillars of physics and gives the modern description of gravitation. It is the basis of our understanding of the evolutionary history and the future of the Universe, and the structure of the Universe at the largest scales. General relativity describes gravity as the curvature of spacetime due to the presence of matter as given by the Einstein equations [1],

$$R_{\mu\nu} - \frac{1}{2} g_{\mu\nu} R = \frac{8\pi G}{c^4} T_{\mu\nu}, \tag{1.1}$$

first written down by Einstein in their final form in 1915 [14]. In this equation, $g_{\mu\nu}$ is the metric tensor while $R_{\mu\nu}$ is the Ricci Tensor and R is the Ricci scalar, both of which are related to the Riemann curvature tensor,

$$R^{\rho}{}_{\sigma\mu\nu} = \partial_\mu \Gamma^{\rho}_{\nu\sigma} - \partial_\nu \Gamma^{\rho}_{\mu\sigma} + \Gamma^{\rho}_{\mu\lambda}\Gamma^{\lambda}_{\nu\sigma} - \Gamma^{\rho}_{\nu\lambda}\Gamma^{\lambda}_{\mu\sigma}, \tag{1.2}$$

that defines the curvature of smooth, differentiable manifolds. Here $\partial_\mu \varphi = \frac{\partial \varphi}{\partial x_\mu}$ and $\Gamma^{\mu}_{\nu\rho}$ are the Christoffel symbols defined in terms of the metric as,

$$\Gamma^{\lambda}_{\mu\nu} = \frac{g^{\lambda\sigma}}{2} \left(\partial_\mu g_{\nu\sigma} + \partial_\nu g_{\mu\sigma} - \partial_\sigma g_{\mu\nu} \right). \tag{1.3}$$

[1] Where ever applicable, the metric signature is chosen to be (-, +, +, +). Furthermore, Greek indices go from 0 to 4 while latin indices go from 1 to 4 and cover only the spatial parts.

The Ricci tensor is defined in terms of the Riemann curvature tensor as $R_{\mu\nu} = R^{\lambda}{}_{\mu\lambda\nu}$, while the Ricci scalar is the trace of the Ricci tensor i.e. $R = g^{\mu\nu}R_{\mu\nu}$.

The term $T_{\mu\nu}$ in Eq. 1.1 is the energy-momentum tensor, while G and c represent Newton's gravitational constant and the speed of light, respectively. The theory of general relativity is invariant under all possible continuous coordinate transformations, and this notion of general covariance is central to the theory.

Equations 1.1 describes the effect of matter on the curvature of spacetime. The curvature, in turn, dictates the motion of matter through the geodesic equation,

$$\frac{d^2x^\mu}{d\lambda^2} + \Gamma^\mu_{\nu\rho}\frac{dx^\nu}{d\lambda}\frac{dx^\rho}{d\lambda} = 0, \tag{1.4}$$

where λ is an affine parameter. One example of such a parameter is the proper time τ.

In the century since its discovery, general relativity has been well validated by an immense breadth of experimental and observational tests. The first success of the theory came as soon as it was developed in explaining the then-mysterious anomalous precession of the perihelion of Mercury's orbit [15]. In 1920, the predicted gravitational lensing of background stars by the Sun was observed during a solar eclipse [16]. The first conclusive evidence of gravitational redshift was observed in the Pound-Rebka experiment in 1960, which measured the frequency shift induced by the Earth's gravitational potential upon the 14.4 KeV γ radiation emitted by Fe^{54} nuclei [17]. This was accompanied shortly by measurement of the gravitational Shapiro time delay in the 1960s, of the round trip travel time of radar pulses from Earth to Venus [18]. These four tests are considered the classical tests of general relativity.

One other key prediction of general relativity is the existence of wave-like solutions linked to the quadrupolar moment of energy and matter. In fact, the existence of such waves had been speculated even earlier with the recognition that a field theory of gravitation is needed in order to make it compatible with special relativity.

Einstein himself first predicted the existence of these waves within the context of general relativity [19]. In his 1918 paper, Einstein recognized that the quadrupole moment provides the leading order contribution to gravitational waves (GWs) instead of the leading order dipole moment contribution for electromagnetic waves [20]. Because of the weakness of gravity and the quadrupolar nature, stellar binary systems were

recognized as relevant sources for GWs. Nevertheless, the nature of GWs and their "realness" was discussed for decades until being finally settled in the affirmative during the Chapel Hill Conference in 1957.

Further debate continued, however, about whether binary inspirals can emit GWs as both stars follow geodesic "inertial" trajectories [2]. A conclusive albeit indirect observational confirmation that binaries do emit gravitational radiation came from the observation of orbital decay of PSR B1913+16 binary pulsar system [21, 22], found to be consistent with the loss of energy and angular momentum through GWs.

The direct detection of GWs had to wait another two and a half decades until the first observing run of the advanced Laser Interferometer Gravitational-Wave Observatory (aLIGO) detectors. On September 14th, 2015, a binary black hole (BBH) merger GW150914 [23] was detected for the first time by the two aLIGO detectors [24] at Hanford, WA and Livingston Parish, LA heralding the beginning of GW astronomy. Then, on August 17th, 2017, GWs from the merger of binary neutron stars (BNS), GW170817, was detected [25] by the aLIGO detectors in conjunction with the advanced Virgo (aVirgo) detector [26] in Italy. Armed with the three-detector localization of the BNS merger, the electromagnetic counterpart was quickly discovered and followed across the electromagnetic spectrum [27], opening up a new chapter in the field of multi-messenger astrophysics. All-in-all, aLIGO-aVirgo have detected about 50 GW events up until the first half of the third observing run (O3).

The ground-based detectors will be complemented soon by the space-based Laser Interferometer Space Antenna (LISA) mission, sensitive to GWs in the millihertz frequency band. LISA has a planned launch date in 2034 and promises the ability to detect GWs from galactic double white dwarf binaries, extreme mass ratio inspirals, supermassive binary black holes, stellar-mass binary black hole inspirals, among other sources.

The rest of the chapter is organized as follows. The physics of GWs in flat spacetime is briefly discussed in Sec. 1.2. Section. 1.3 describes the interaction of GWs with matter in flat spacetime which is used to calculate the coupling of GWs to interferometric detectors in Sec. 1.4. The astrophysical and cosmological sources of GWs are discussed in Sec. 1.5 while Sec. 1.6 describe some of the instrumental details and important noise

[2] This was the gravitational analog of the question of whether a freely falling charge radiates.

sources of ground-based detectors. We will finish the chapter with a brief discussion about GW astrophysics done thus far in Sec. 1.7.

1.2 Gravitational-waves

This section will briefly discuss some relevant physics of GWs, focusing in particular on understanding their coupling with detectors. It will follow closely Chapter 1 of Ref. [28] to which the reader is directed for more details. Outside of a source, the energy-momentum tensor is zero, and the Einstein equations reduce to,

$$R_{\mu\nu} - \frac{1}{2} g_{\mu\nu} R = 0. \tag{1.5}$$

Far away from the source, GWs can be best understood using a linearized theory where the metric $g_{\mu\nu}$ can be expanded around flat spacetime as,

$$g_{\mu\nu} \approx \eta_{\mu\nu} + h_{\mu\nu}. \tag{1.6}$$

where $h_{\mu\nu} \ll 1$ and $\eta_{\mu\nu}$ is the Minkowski metric:

$$\eta_{\mu\nu} = \begin{pmatrix} -1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \tag{1.7}$$

that defines the structure of flat spacetime in special relativity. The linearized theory is now only invariant under infinitesimal local transformations $x^\mu \to x^\mu + \xi^\mu$, where ξ^μ is similarly small. Nevertheless, a high degree of coordinate symmetries remain, to be fixed by making appropriate gauge choices. Let's define the trace of $h_{\mu\nu}$ as $h = \eta^{\mu\nu} h_{\mu\nu}(x)$ and let,

$$\bar{h}_{\mu\nu} = h_{\mu\nu} - \frac{1}{2} \eta_{\mu\nu} h. \tag{1.8}$$

The Einstein equations in linearized theory reduce to [28],

$$\Box\bar{h}_{\mu\nu} + \eta_{\mu\nu}\partial^\rho\partial^\sigma\bar{h}_{\rho\sigma} - \partial^\rho\partial_\nu\bar{h}_{\mu\rho} - \partial^\rho\partial_\mu\bar{h}_{\nu\rho} = 0. \tag{1.9}$$

We now choose the *Loretnz gauge* or the *De Donder* gauge defined by the condition,

$$\partial^\mu\bar{h}_{\mu\nu} = 0, \tag{1.10}$$

which decreases the number of independent degrees of freedom to 6 and reduces the Einstein equations to,

$$\Box\bar{h}_{\mu\nu} = 0. \tag{1.11}$$

But an additional gauge choice can be made by choosing the transverse-traceless (TT) gauge which effectively amounts to the condition that $\Box\xi^\mu = 0$. In the TT gauge, we choose to set the trace h and all components h^{0i} to be zero leaving us with two degrees of freedom. We can then find plane-wave solutions to Eq 1.11 with a direction of propagation $\hat{\mathbf{n}}$ [3]:

$$h_{ij}^{\mathrm{TT}}(t,\mathbf{x}) = \sum_{A=+,\times}\int_{-\infty}^{\infty} df\,\tilde{h}_A(f)\,e_{ij}^A(\hat{\mathbf{n}})\exp\left\{-2\pi i f(t-\hat{\mathbf{n}}\cdot\mathbf{x}/c)\right\}, \tag{1.12}$$

where $i,j \in \{1,2,3\}$, $A = +,\times$ are the two polarization states and $\tilde{h}_A^*(f)$ is the Fourier component of a GW of frequency f and polarization A. The polarization tensors e_{ij}^A are defined in terms of two units vectors $\hat{\mathbf{p}},\hat{\mathbf{q}}$ orthogonal to $\hat{\mathbf{n}}$:

$$e_{ij}^+(\hat{\mathbf{n}}) = \hat{\mathbf{p}}_i\hat{\mathbf{p}}_j - \hat{\mathbf{q}}_i\hat{\mathbf{q}}_j, \qquad e_{ij}^\times(\hat{\mathbf{n}}) = \hat{\mathbf{p}}_i\hat{\mathbf{q}}_j + \hat{\mathbf{q}}_i\hat{\mathbf{p}}_j. \tag{1.13}$$

The reality of $h_{ij}^{\mathrm{TT}}(t,\mathbf{x})$ requires that $\tilde{h}_A(-f) = \tilde{h}_A(f)$. The TT reference frame expression extends quite easily to the case where there are more than one GWs acting at $(t,\mathbf{x})$ and in the most general case can be written as,

$$\boxed{h_{ij}^{\mathrm{TT}}(t,\mathbf{x}) = \sum_{A=+,\times}\int_{-\infty}^{\infty} df\int_{4\pi} d^2\hat{\mathbf{n}}\,\tilde{h}_A(f,\hat{\mathbf{n}})\,e_{ij}^A(\hat{\mathbf{n}})\exp\left\{-2\pi i f(t-\hat{\mathbf{n}}\cdot\mathbf{x}/c)\right\}.} \tag{1.14}$$

[3] In this dissertation bold fonts should be taken to mean three-dimensional vectors unless specified otherwise.

This expression gives us the most generic wave solution to Einstein's equations in flat spacetime.

1.3 Coupling of GWs with Matter

In order to understand how interferometric detectors work, it is necessary to understand how GWs affect the world lines of particles, especially freely falling ones. This section calculates the coordinate separation and the proper distance between spatially separated test masses in a GW's presence, adopting the convenient TT frame.

1.3.1 Coordinate separation and proper distance

The effect of GWs on coordinate separation is best studied by studying the geodesic equation, Eq. 1.4. Let x^μ be the geodesic of a test mass which is initially at rest in the TT frame as a GW passes by. The spatial components of the geodesic change as,

$$\frac{d^2x^i}{d\tau^2} = -\Gamma^i_{\nu\rho}\frac{dx^\nu}{d\tau}\frac{dx^\rho}{d\tau} \tag{1.15}$$

At $\tau = 0$ since the mass was at rest, we get,

$$\left.\frac{d^2x^i}{d\tau^2}\right|_{\tau=0} = -\left.\Gamma^i_{00}\left(\frac{dx^0}{d\tau}\right)^2\right|_{\tau=0}. \tag{1.16}$$

In the linearized TT gauge, the Christoffel symbol Γ^i_{00} becomes,

$$\Gamma^i_{00} \approx \frac{\eta^{ii}}{2}\left(\partial_0 h_{0i} + \partial_0 h_{0i} - \partial_i h_{00}\right). \tag{1.17}$$

But as discussed before, the components h_{0i} are zero in the TT frame. The component h_{00} corresponds to the static gravitational field and can also be set to zero here. Hence we have,

$$\left.\frac{d^2x^i}{d\tau^2}\right|_{\tau=0} = 0 \tag{1.18}$$

This implies that in the TT frame, an inertial test mass initially at rest remains at rest as a GW passes [28]. A corollary is that the coordinate separation between two freely

falling test masses initially at rest does not change with time in the TT gauge [4]!

Nevertheless, while the coordinate separation doesn't change, a passing GW does have an effect on the proper distance between spacetime events in the TT frame. Consider a light ray which travels from test mass 1 to test mass 2, corresponding to events $(t_1, \mathbf{x}_1)$ and $(t_2, \mathbf{x}_2)$ respectively. The two masses are at rest and their spatial coordinate separation is given by $L_0 = |\mathbf{x}_1 - \mathbf{x}_2|$. The spacetime interval under the linear approximation is,

$$ds^2 = g_{\mu\nu} dx^\mu dx^\nu = (\eta_{\mu\nu} + h_{\mu\nu}) dx^\mu dx^\nu = 0. \tag{1.19}$$

In the TT frame, this becomes

$$ds^2 = -cdt^2 + (\delta_{ij} + h_{ij}(t, \mathbf{x}))\, dx^i dx^j = 0. \tag{1.20}$$

We define the Euclidean distance element, $dl_0^2 = \delta_{ij} dx^i dx^j$. The proper distance element is thus,

$$dl^2 = dl_0^2 + h_{ij}(t, \mathbf{x}) dx^i dx^j \implies \left(\frac{dl}{dl_0}\right)^2 = 1 + h_{ij}(t, \mathbf{x}) \frac{dx^i}{dl_0} \frac{dx^j}{dl_0}. \tag{1.21}$$

Under the linear approximation, $h_{ij} \ll 1$, we get,

$$\frac{dl}{dl_0} = \sqrt{1 + h_{ij}(t, \mathbf{x}) \frac{dx^i}{dl_0} \frac{dx^j}{dl_0}} \approx 1 + \frac{1}{2} h_{ij}(t, \mathbf{x}) \frac{dx^i}{dl_0} \frac{dx^j}{dl_0}. \tag{1.22}$$

The proper distance between the two events can then be obtained by integration:

$$l = L_0 + \frac{1}{2} \int_0^{L_0} dl_0\, h_{ij}(t, \mathbf{x}) \frac{dx^i}{dl_0} \frac{dx^j}{dl_0}. \tag{1.23}$$

To see a concrete example that the proper distance is measurable, let us take the simplistic case under which there is a single plane GW of frequency f from the direction of $\hat{\mathbf{z}}$. Furthermore, let us assume the coordinate distance between the two masses be much smaller than the GW wavelength – this is a good approximation for the aLIGO-aVirgo detectors – and that they lie along the x-axis. Using the expression for the GW

[4] One can, of course, construct other frames where this is not true, and the coordinate separation does change. One example is the so-called detector frame which conforms more readily to our Newtonian intuitions. The reader is again pointed to Ref. [28] for a more detailed discussion.

metric from Eq. 1.12, the proper distance in Eq. 1.23 reduces to,

$$l(t) = l_0 \left(1 + h_+ \cos(2\pi f t)\right). \tag{1.24}$$

This shows the proper distance changes with time and, in fact, oscillates with the same frequency as the passing GW.

1.4 Laser Interferometry

As described in the previous section, the proper distance between events provides a good metric to detect a passing GW. This motivates the use of interferometry in general and laser interferometry in particular to detect GWs. This section will calculate the coupling of GWs with a Michelson laser interferometer.

1.4.1 Coupling of GW to Michelson Interferometers

A Michelson interferometer measures the differential signal across two arms of the round-trip travel of the laser light. A simplified schematic of such an interferometer is shown in Fig. 1.1. Let us suppose one of the arms of the Michelson interferometer of coordinate length L_0 is along the vector $\hat{\mathbf{u}}$. We will assume that the interferometer can be approximated to be stationary in the time period of interest. Without loss of generality, we assume that the beamsplitter is at the origin of the coordinate system. Let us consider an incoming GW signal from the direction of $\hat{\mathbf{n}}$. Exploiting the Fourier transform's linearity, we work in the frequency basis assuming a GW of frequency f. A beam of light leaves the beam splitter at time t_0, bounces off the end mirror at coordinate time $t_1 = L_0/c$ and reaches back the beam splitter at time $t_2 = 2L_0/c$. This marks off the spacetime events $e_1 = (0, \mathbf{0})$, $e_2 = (L_0/c, L_0\hat{\mathbf{u}})$ and $e_3 = (2L_0/c, \mathbf{0})$. Let us first consider the spacetime interval in the TT frame between the first two events following Eq. 1.23,

$$l_{12}(f) = L_0 + \frac{1}{2} \sum_{A=+,\times} \int_0^{L_0} dl\, \tilde{h}_A(f) \exp\left(-\frac{2\pi i f l}{c}\left(1 - \hat{\mathbf{n}}\cdot\hat{\mathbf{u}}\right)\right) e^A_{ij}(\hat{\mathbf{n}})\, u^i u^j, \tag{1.25}$$

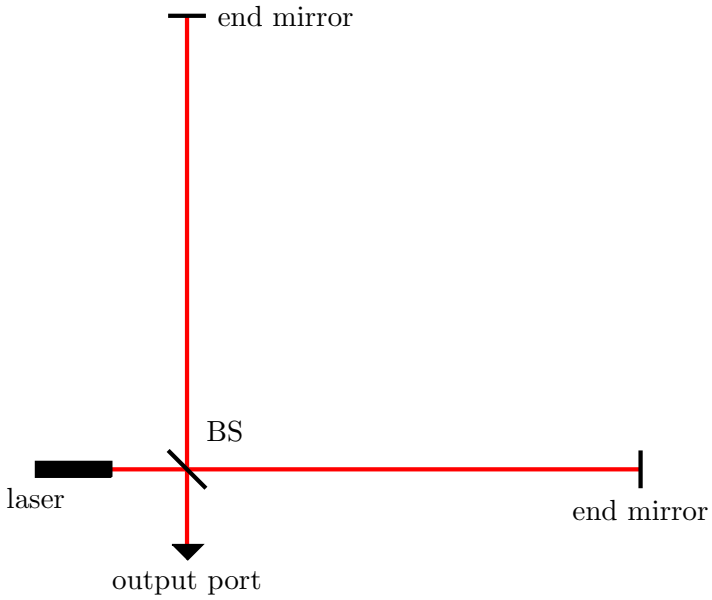


Figure 1.1: A schematic of a Michelson interferometer. The beam splitter, BS, divides the laser light from the laser, which reflects off of the end mirrors in each arm and interferes back at the center. Because the change in the light's phase depends on the proper distance, a passing GW can change this interference, as seen at the output port. Note that while the figure shows perpendicular arms, this is not necessary, and detectors like LISA have an angle of 60° between the arms.

where we have used the fact that the coordinate time $t = l/c$ for the laser beam. Upon integrating we get,

$$l_{23}(f) = L_0 + \frac{L_0}{2} \sum_{A=+,\times} u^i u^j e^A_{ij}(\hat{\mathbf{n}})\, \tilde{h}_A(f)\, \mathrm{sinc}\left(\frac{\pi f L_0}{c}\,(1 - \hat{\mathbf{n}}\cdot\hat{\mathbf{u}})\right) \times \exp\left(-\frac{i\pi f L_0}{c}\,(1 - \hat{\mathbf{n}}\cdot\hat{\mathbf{u}})\right), \tag{1.26}$$

where $\mathrm{sinc}\, x = \sin x/x$. Similarly, the expression for the back trip corresponding to the events e_2 and e_3 is,

$$l_{12}(f) = L_0 + \frac{L_0}{2} \sum_{A=+,\times} u^i u^j e^A_{ij}(\hat{\mathbf{n}})\, \tilde{h}_A(f)\, \mathrm{sinc}\left(\frac{\pi f L_0}{c}\,(1+\hat{\mathbf{n}}\cdot\hat{\mathbf{u}})\right) \times \exp\left(-\frac{i\pi f L_0}{c}\,(3+\hat{\mathbf{n}}\cdot\hat{\mathbf{u}})\right) \tag{1.27}$$

Putting the two together for the roundtrip calculation and defining the strain as $\tilde{h}_u(f) = (l_{12} + l_{23} - 2L_0)/2L_0$, we get,

$$\tilde{h}_u(f) = \frac{1}{2}\mathcal{T}(f, \hat{\mathbf{u}}\cdot\hat{\mathbf{n}})\, u^i u^j \sum_{A=+,\times} \tilde{h}_A(f)\, e^A_{ij} \tag{1.28}$$

where we have defined the round-trip transfer function $\mathcal{T}(f, \hat{\mathbf{u}}\cdot\hat{\mathbf{n}})$ as,

$$\mathcal{T}(f, \hat{\mathbf{u}}\cdot\hat{\mathbf{n}}) = \frac{1}{2}\Bigg[\mathrm{sinc}\left(\frac{\pi f L_0}{c}\,(1-\hat{\mathbf{n}}\cdot\hat{\mathbf{u}})\right)\exp\left(-\frac{i\pi f L_0}{c}\,(1-\hat{\mathbf{n}}\cdot\hat{\mathbf{u}})\right) + \mathrm{sinc}\left(\frac{\pi f L_0}{c}\,(1+\hat{\mathbf{n}}\cdot\hat{\mathbf{u}})\right)\exp\left(-\frac{i\pi f L_0}{c}\,(3+\hat{\mathbf{n}}\cdot\hat{\mathbf{u}})\right)\Bigg]. \tag{1.29}$$

Similarly, the strain observed in the other arm along the vector $\hat{\mathbf{v}}$ is given by,

$$\tilde{h}_v(f) = \frac{1}{2}\mathcal{T}(f, \hat{\mathbf{v}}\cdot\hat{\mathbf{n}})\, v^i v^j \sum_{A=+,\times} \tilde{h}_A(f)\, e^A_{ij} \tag{1.30}$$

The Michelson signal is then defined as the differential strain, i.e. $h_u(f) - h_v(f)$:[5],

$$\tilde{h}_{\mathrm{mich}}(f, \hat{\mathbf{n}}) = \frac{1}{2}\left[\mathcal{T}(f, \hat{\mathbf{u}}\cdot\hat{\mathbf{n}})\, u^i u^j - \mathcal{T}(f, \hat{\mathbf{v}}\cdot\hat{\mathbf{n}})\, v^i v^j\right] \sum_{A=+,\times} \tilde{h}_A(f)\, e^A_{ij}. \tag{1.31}$$

thus yielding the general frequency-domain expression for the response of a laser interferometric detector to GWs with arms along $\hat{\mathbf{u}}$ and $\hat{\mathbf{v}}$. This is consistent with the doppler tracking derivation presented in literature; see, for example, Refs. [29, 30, 31]. The time-domain expression, if necessary, can simply be obtained by inverse Fourier

[5] Note that while all constructed or proposed interferometric GW detectors have arms of equal length, this is by no means a necessity.

transformation. It is convenient to define the antenna pattern functions $F^A(f)$ for individual polarization modes,

$$F^A(f,\hat{\mathbf{n}}) = \frac{1}{2}\left[\mathcal{T}(f,\hat{\mathbf{u}}\cdot\hat{\mathbf{n}})\,u^i u^j - \mathcal{T}(f,\hat{\mathbf{v}}\cdot\hat{\mathbf{n}})\,v^i v^j\right] e^A_{ij}. \tag{1.32}$$

In terms of the antenna pattern functions, the Michelson signal can be written down as,

$$\tilde{h}_{\rm mich}(f,\hat{\mathbf{n}}) = \tilde{h}_+(f)F^+(f) + \tilde{h}_\times(f)F^\times(f). \tag{1.33}$$

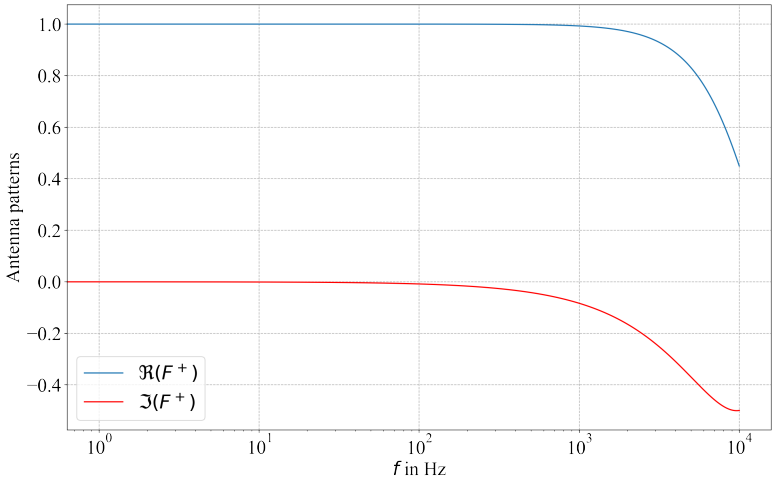
Figure 1.2 shows an example of the frequency dependence of Michelson antenna patterns for two different detectors to a GW traveling perpendicular to the plane defined by the arms. In the case of an aLIGO-like detector with 4 km long arms, the antenna patterns are constant for most of the range and only start to change above $\sim$ 1000 Hz significantly. As we shall see in the Sec. 1.6, aLIGO and aVirgo are bound by various types of noise at frequencies less than 10 Hz and frequencies greater than 1000 Hz. In the "bucket" between these frequencies, the detector's response to a GW is well approximated to be independent of frequency,

$$F^A(\hat{\mathbf{n}}) \approx \frac{1}{2}\left(u^i u^j - v^i v^j\right) e^A_{ij}. \tag{1.34}$$

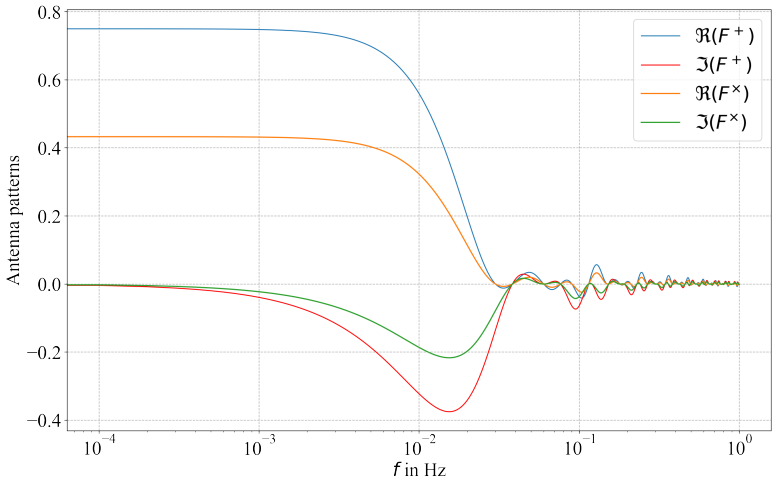
and aLIGO is thus said to function in the large-wavelength limit where $\lambda_{GW} \gg L_0$. One can thus write Eq. 1.33 in the time domain for LIGO-Virgo as,

$$h_{\rm mich}(t) = h_+(t)F^+(t,\hat{\mathbf{n}}) + h_\times(t)F^\times(t,\hat{\mathbf{n}}) \tag{1.35}$$

The situation is very different for a LISA which has an arm length of 2.5 million km, as shown by the antenna patterns in Fig. 1.2. While the antenna patterns are constants at low frequencies, they show strong frequency dependence between $10^{-3} - 3\times10^{-2}$ Hz and oscillate closely around zero above that range. Thus, while in the millihertz range $\lambda_{GW} \simeq L_0$, LISA operates in the small-wavelength limit above 3×10^{-2} Hz where $\lambda_{GW} \ll L_0$. In fact, unlike the aLIGO-aVirgo detectors, the sensitivity of LISA is limited at high frequencies not by sources of noise but by the physics of GWs itself because of the small values of the antenna patterns!



(a)



(b)

Figure 1.2: The top figure plots the antenna patterns for an aLIGO-like detector with an arm length of 4 km towards a GW traveling along the z axis. The arms are oriented at right angles with respect to each other, along the coordinate axes in the $x-y$ plane, The antenna pattern for $\times$ polarization vanishes uniformly because of the symmetry of the system. The bottom figure plots the antenna patterns for a LISA-like Michelson detector with 2.5 million km long arms, one of which is oriented along the x axis. The $\times$ polarization does not vanish in this case because the angle between the arms is 60°.

1.5 Sources of Gravitational Waves

In order to understand the types of sources a GW detector could detect, it is helpful to calculate the power carried by GWs. We work with linearized theory again under weak field and low-velocity approximations, using a multipole expansion. The leading term in the multipole expansion is the quadrupole moment terms [6]. The GW metric from the quadrupole term in the TT-gauge at a point $\mathbf{x}$ far away from the source is given by,

$$h_{ij}^{\mathrm{TT}}(t,\mathbf{x}) = \frac{2G}{rc^4}\ddot{Q}_{ij}^{\mathrm{TT}}(t-r/c), \tag{1.36}$$

where $r = |\mathbf{x}|$ and Q_{ij}^{TT} is the quadrupole moment tensor:

$$Q_{ij}^{\mathrm{TT}} = \frac{1}{c^2}\int d^3x\, T^{00}\left(x_i x_j - \frac{x_k x_k}{3}\delta_{ij}\right). \tag{1.37}$$

While GWs are generated by mass and energy as attested by Eq. 1.36, they are themselves sources of energy and momentum too. Under linearized theory, we can describe the energy-momentum tensor of GWs as,

$$t_{\mu\nu} = \frac{c^4}{32\pi G}\langle \partial_\mu h_{\alpha\beta}\partial_\nu h^{\alpha\beta}\rangle, \tag{1.38}$$

where $\langle ...\rangle$ indicates a spatiotemporal average over scales much longer than the GW wavelength and much smaller than the GW frequency. In the TT gauge this equation becomes,

$$t_{\mu\nu} = \frac{c^4}{32\pi G}\langle \partial_\mu h_{ij}^{\mathrm{TT}}\partial_\nu h_{ij}^{\mathrm{TT}}\rangle. \tag{1.39}$$

In particular, the energy density is given by,

$$t_{00} = \frac{c^2}{32\pi G}\langle \dot{h}_{ij}^{\mathrm{TT}}\dot{h}_{ij}^{\mathrm{TT}}\rangle = \frac{c^2}{16\pi G}\langle \dot{h}_+^2 + \dot{h}_\times^2\rangle. \tag{1.40}$$

We derive an expression for the energy flux by integration:

[6] In addition to the well-known weakness of gravity compared to other fundamental forces, the fact that the leading order term is quadrupole - rather than dipole as in electromagnetism - also contributes to the weakness of GWs.

$$\frac{dE}{dt} = \frac{c^3 r^2}{16\pi G} \int d\Omega \langle \dot{h}_+^2 + \dot{h}_\times^2 \rangle. \tag{1.41}$$

Combining this with Eq. 1.36 gives an expression for the GW power in the quadrupole approximation as,

$$P_{\rm quad} = \frac{dE}{dt} = \frac{G}{5c^5} \langle \ddot{Q}_{ij}^{\rm TT} \ddot{Q}_{ij}^{\rm TT} \rangle. \tag{1.42}$$

Finally, we can also write the expression for the energy flux in Fourier domain which is often more convenient for purposes of data analysis as,

$$\frac{dE}{df} = \frac{\pi c^3 f^2 r^2}{2G} \int d\Omega \left\{ |\tilde{h}_+(f)|^2 + |\tilde{h}_\times(f)|^2 \right\}. \tag{1.43}$$

1.5.1 Compact binary coalescence

Due to factors such as the leading order quadrupolar term, the weakness of gravity and the high velocities needed for relativistic effects to kick in, compact binary coalescences (CBC) are one of the most important sources of GWs. This includes stellar-mass BBH and BNS mergers in the aLIGO frequency band and LISA sources like double-white dwarfs (DWD) binaries and extreme mass ratio inspirals (EMRIs).

Let m_1 and m_2 be two such objects in a binary orbit at a distance d from us. Under the assumption that the orbits are quasi-circular, one can solve Eq. 1.36 to get the quadrupolar GW strain from the binary,

$$\begin{aligned} h_+(t) &= \frac{4}{d} \left(\frac{G\mathcal{M}}{c^2} \right)^{5/3} \left(\frac{\pi f_{gw}(t)}{c} \right)^{2/3} \left(\frac{1 + \cos^2 \iota}{2} \right) \cos\left(2\pi f_{gw} t + \varphi_0\right), \\ h_\times(t) &= \frac{4}{d} \left(\frac{G\mathcal{M}}{c^2} \right)^{5/3} \left(\frac{\pi f_{gw}(t)}{c} \right)^{2/3} \cos\iota \sin\left(2\pi f_{gw} t + \varphi_0\right) \end{aligned} \tag{1.44}$$

Here, $\mathcal{M}$ is the chirp mass defined as,

$$\mathcal{M} = \frac{(m_1 m_2)^{3/5}}{(m_1 + m_2)^{1/5}}, \tag{1.45}$$

and the angle ι defines the inclination between the line of sight to the binary and the normal to the plane of the binary. The GW frequency f_{gw} in Eq 1.44 is twice the orbital

frequency and evolves with time as,

$$f_{gw}^{-8/3}(t) = \frac{(8\pi)^{8/3}}{5}\left(\frac{G\mathcal{M}}{c^3}\right)^{5/3}(t_{\rm col} - t), \tag{1.46}$$

where $t_{\rm col}$ is the coalescence time of the binary.

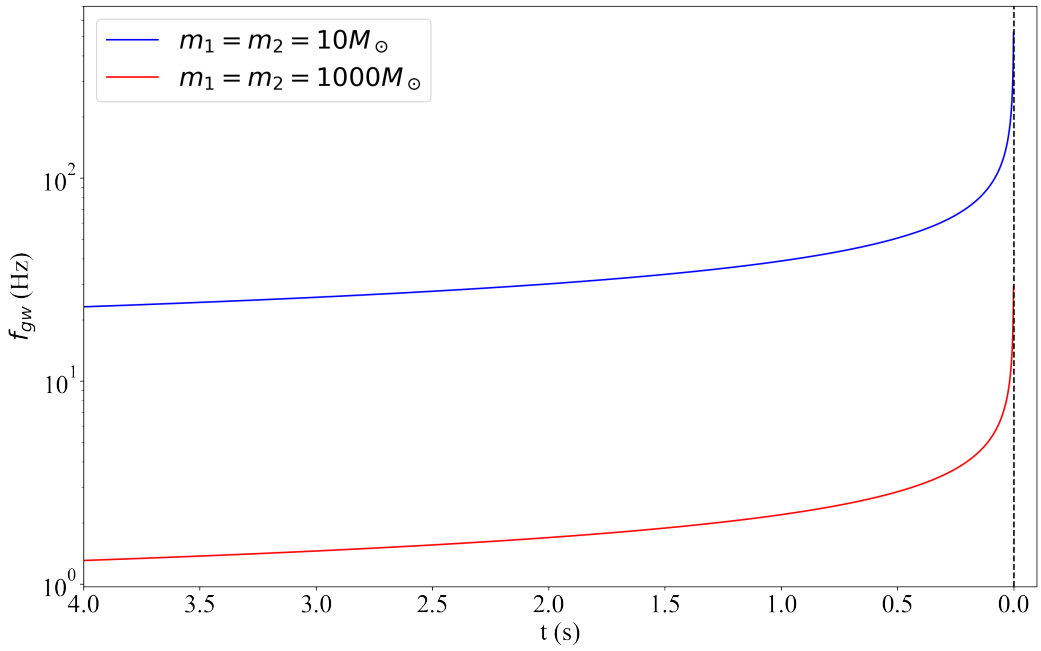


Figure 1.3: The evolution of the GW frequency of two binary systems is shown assuming the quadrupolar approximation. Both binaries coalesce at $t_{\rm col} = 0$, represented by the dashed black line, at which the GW frequency under this approximation diverges.

Figure. 1.3 shows the frequency evolution of two binary systems under the quadrupolar approximation. The masses of both objects in the first binary is $10\,M_\odot (\mathcal{M} \approx 8.70\,M_\odot)$ and $1000\,M_\odot (\mathcal{M} \approx 870\,M_\odot)$ for the second binary. In the last few seconds before a merger, where the metric perturbations are the strongest, the first binary spends much more of its time in the sensitive frequency band between $\sim 10 - 1000$ Hz compared to the second binary. This is the main reason why aLIGO and aVirgo are sensitive to stellar-mass binaries but not high mass intermediate-mass binaries or supermassive

binaries.

Consider the $m_1 = m_2 = 10M_\odot$ binary optimally oriented towards us i.e $\iota = 0$ at a distance of 100 Mpc. Using Eqs. 1.44 one can estimate that the strain from the binary around a frequency of 100 Hz will be about $h \sim 10^{-21}$. In other words, we need the noise floor to be lower than 10^{-21} to be able to detect GWs. Historically this was the target which was aimed for in the development of the aLIGO-aVirgo detectors.

Finally, while linearized multipole expansion is adequate for approximately calculating the energy and momentum carried away by GWs, it fails to capture the strong field dynamics of CBC sources. This is especially important in the last few moments of the inspiral during which such strong field dynamics dominate. Developments in post-Newtonian theory and numerical relativity [32, 33, 34, 35] in the past few decades have been instrumental in this regard, enabling us to construct good waveform models to capture the dynamical features of CBCs adequately. These waveform models allow building matched-filter detection algorithms for CBC sources that represent the mathematically optimal detection method in the presence of Gaussian instrumental noise. These sophisticated waveforms have played a predominant role in most detections made thus far by aLIGO-aVirgo and in Bayesian parameter estimation algorithms for CBC sources. The latter allows us to accurately determine the physical properties of the black holes and neutron stars sources from the GW data.

1.5.2 Gravitational-wave bursts

Unlike CBC signals, bursts constitute a wide range of transient GW sources whose signal morphologies are not well understood. GWs from core-collapse supernova constitute an important source of such signals [36]. Other sources include rotational glitches in neutron stars [37], GWs from accretion disk instabilities [38, 39], magnetar bursts and flares [40, 41], fallback accretion onto new neutron stars [41], etc. Bursts are also possible from more speculative sources like cosmic string cusps [42, 43].

Since the waveform features of GW burst are unknown, they are usually targeted by unmodeled excess-power searches. Such algorithms do not make detailed waveform assumptions but rather attempt to detect the excess power of the GWs in multiple detectors. Chapter 3 will discuss one such algorithm in specific detail. Other burst algorithms decompose the GW signal on some universal basis, like a wavelet basis coherently over

multiple detectors. Note that while these algorithms make no modeling assumptions, they can work perfectly well for well-modeled CBC sources too, and present a critical cross-check of waveform modeling itself.

1.5.3 Continuous-wave sources

Continuous wave or periodic sources are long-lived sources of GWs which emit nearly monochromatic radiation over long periods. We can track such a source for a long time – months or even years – allowing us to detect much fainter GWs than is possible with transients. Modeling these sources is also relatively more straightforward as they show little frequency evolution over these time scales. Nevertheless, the need to integrate for such a long time with fine frequency resolution introduces unique computational and data quality challenges to continuous-wave algorithms.

In the aLIGO-aVirgo frequency band, the main source of periodic GWs are rapidly spinning neutron stars with a deformation [44], usually represented with the ellipticity parameter ϵ. This induces a non-zero quadrupolar moment allowing it to radiate GWs. If the neutron star is spinning along the z axis and I_x, I_y and I_z are the moments of inertia along the x, y and z axes, then the ellipticity is given by,

$$\epsilon = \frac{|I_x - I_y|}{I_z}. \tag{1.47}$$

In the LISA band, double white dwarf binaries, in which the two white dwarfs are widely separated, can be a continuous-wave source. In either case, accurate modeling will also have to account for doppler modulations caused by the Earth's orbit and potentially also of the source.

1.5.4 Stochastic gravitational-wave backgrounds

The final category of GWs to be considered here are stochastic gravitational-wave backgrounds (SGWBs). These are broadband GWs, defined only through their statistical properties. One source of these waves could be a primordial GW background generated through inflationary mechanisms [45, 46, 47], analogous to the cosmic microwave background. The discovery of this hypothetical inflationary background remains a primary goal of current and future ground-based GW detectors. A more significant source of

SGWB for aLIGO-aVirgo comes from an incoherent superposition of individually undetectable CBC signals [48, 49, 50, 51]. Other potential sources of SGWBs can also similarly arise from an incoherent superposition of GWs from supernovae [52, 53] or neutron stars [54, 49]. More speculative SGWBs are also possible from cosmic strings [43, 55] or early universe phase transitions (see for e.g [56, 57]). One can also get SGWBs from an incoherent superposition of intermediate mass binary black holes and supermassive binary black holes in the LISA band and the pulsar timing array band [58]. There is also a stochastic galactic foreground that arises from the superposition of galactic double white dwarfs [59]. The foreground is so named because it peaks above the instrumental noise floor of LISA. Section

SGWB detection algorithms attempt to measure the statistical properties that define it. Generally, the SGWB is assumed to be Gaussian and stationary, which allows us to characterize it with its variance completely; or equivalently its power spectrum in the frequency domain $S_{gw}(f)$. Conventionally, the power spectrum is written in terms of the dimensionless energy density per logarithmic frequency bin $\Omega_{gw}(f)$ by applying Eq. 1.43,

$$\Omega_{gw}(f) = \frac{2\pi^2 f^3}{3H_0^2} S_{gw}(f) \tag{1.48}$$

Chapter 6 will discuss stochastic backgrounds and foreground in more detail in the context of LISA.

From the perspective of detection, the four categories of GWs signals are often described based on their temporal nature (transient or persistent) and our understanding of the waveform. This division is schematically shown in Fig. 1.4. However, the division is more continuous; for instance, there can be transient signals that last thousands of seconds or even longer. Chapter 3 discusses an algorithm targeting such signals in more detail.

1.6 Ground-based Laser Interferometric Detectors

This section will focus on the some of the particular details of ground-based GW detectors, in particular the aLIGO and aVirgo detectors and describe some of the key sources of noise which affect the sensitivity of the detectors.

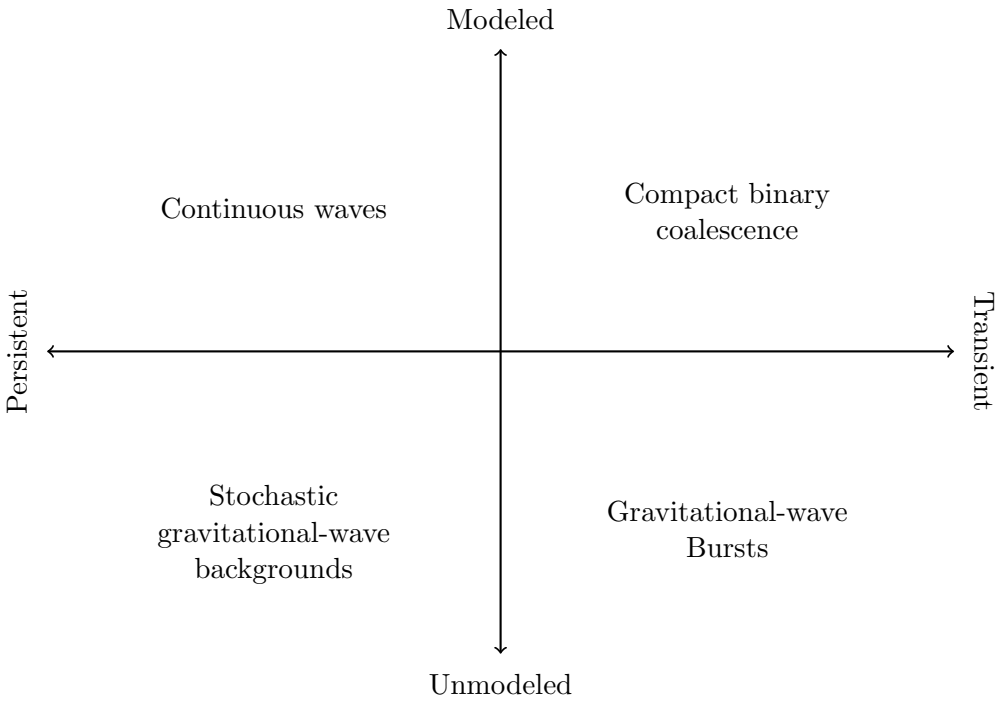


Figure 1.4: The four-fold categorization of GW sources is shown from the point of view of detection algorithms.

As of the writing of this dissertation, there are three GW detectors with sensitivities sufficient for astrophysical detection. They are the two aLIGO detectors with 4 km arms in the USA and the aVirgo detector with 3 km arms in Italy. Additionally, the new KAGRA detector with 3 km arms [60] briefly came online at the end of the third observing run. GEO600 [61] is a European detector with 600 m long arms that is often used for prototyping and testing new detector technologies. A planned detector with 4 km arms is also expected to come online in India [62, 63] in the second half of the 2020s. A map of the detectors is shown in Fig. 1.5.

1.6.1 The LIGO and Virgo interferometric detectors

The aLIGO and aVirgo detectors are Fabry-Pérot cavity-based Michelson interferometers. The resonant Fabry-Pérot cavities are within two perpendicular arms, which are 4 km long in the case of aLIGO and 3 km long in the case of aVirgo, and boost the

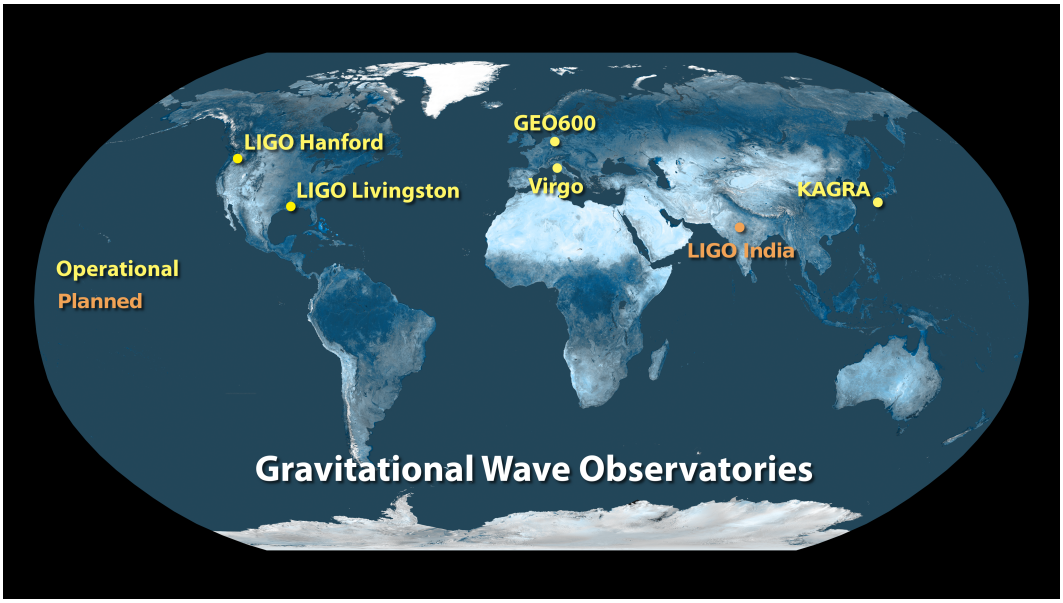


Figure 1.5: A map of all operational and planned ground-based GW observatories. Image credits: Lynn Cominsky and Caltech press office.

interference differential phase at the output port of the interferometer in response to a passing GW. The default state of the interferometer is to have destructive interference so the laser power at the output zero [7]. There are two additional resonant cavities; a power-recycling cavity which increases the laser power in the interferometer by recycling all the light which exits towards the laser, and a signal recycling mirror to improve the broadband frequency response of the Fabry-Pérot cavities. These cavities' lengths are fixed using the Pound-Drever-Hall technique, using radio frequency sidebands superimposed onto 1064 nm laser light. The laser frequency is stabilized to the common arm cavity, which is the sum of the two cavities. Beginning with O3, there is also an in-vacuum squeezer designed to inject squeezed vacuum states into the anti-symmetric port and reduce the electromagnetic quantum noise. A schematic of the optical layout of the aLIGO detectors is shown in Fig. 1.6.

There are four test masses at the end of the Fabry-Pérot cavities, which also serve as mirrors for the reflecting laser light. These test masses and the central beam splitter,

[7] However, a tiny offset is introduced from perfect destructive interference to be sensitive to GWs, see [28]

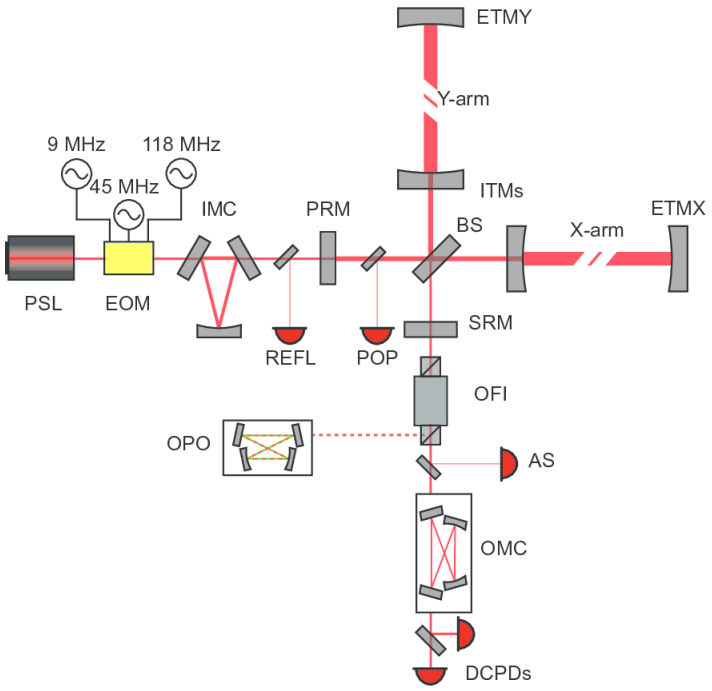


Figure 1.6: A schematic of the optical layout of the LIGO detectors as of O3 is shown. The figure is reproduced from Ref. [4] with permission.

along with the power and signal recycling mirrors are suspended on quadruple pendula to mitigate seismic noise in the detector. The detectors are said to be locked when the laser light is fully resonant in each of the interferometer cavities, with the laser frequency stabilized. The lock is achieved and maintained via a sophisticated sensing and control system using the Pound-Drever-Hall technique [64]. A detailed review of all the methods and techniques that make a functioning Fabry-Pérot interferometer is beyond the scope of this dissertation and the reader is directed to Refs. [4, 26, 24, 65] and the references therein for more details.

1.6.2 Noise sources

For the rest of this section, we describe some of the primary noise sources relevant to the ground-based GW detectors. These noise sources define the sensitivity and the background of any search algorithms. As we saw in Sec.1.5, it was realized very early into the history of development of GW detectors that a minimum strain sensitivity of

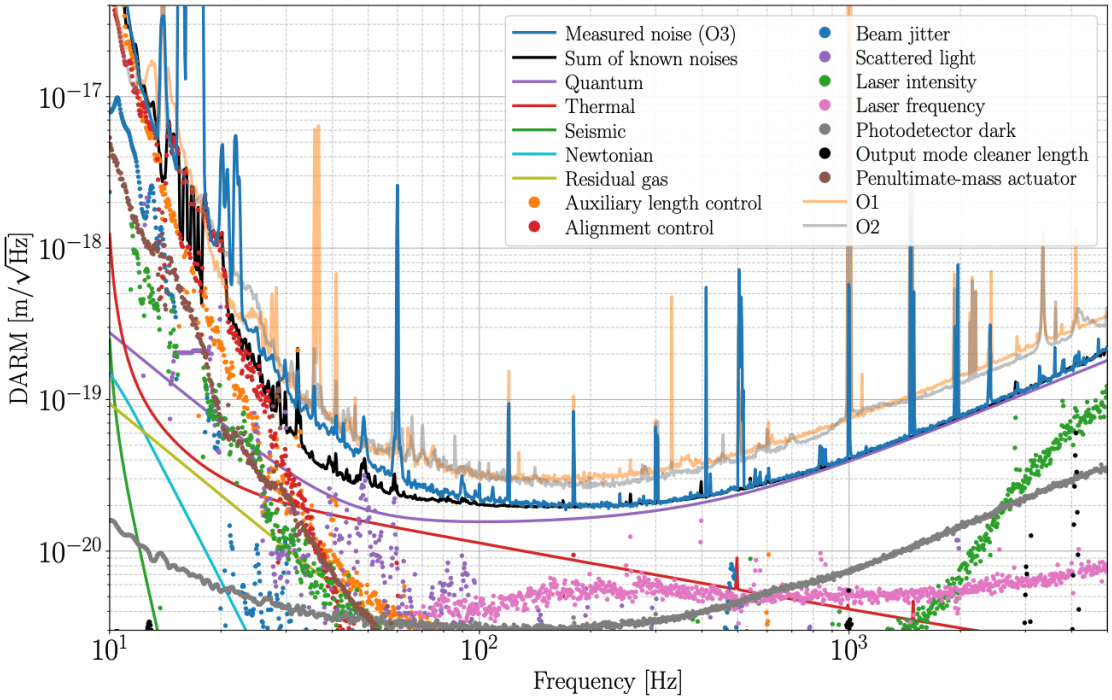


Figure 1.7: The amplitude spectral density of the noise in the aLIGO Hanford detector as a function of frequency is shown. This budget plots also plots the estimates of the various noise sources. The sensitivity is bounded mainly by thermal and seismic noise at low frequencies and by quantum noise at high frequencies. Reproduced from Ref. [4] with permission.

10^{-21} was necessary for detecting GWs in the audible frequency band. Reducing the coupling of noise sources has thus been significant work, both on the original and the advanced detectors. Several key proposed advancements for next-generation detectors like Cosmic Explorer and the Einstein Telescope also target reduction in various noise source couplings. Figure 1.7 shows a noise budget plot with the measured noise levels in the aLIGO Hanford detector during the O3 run, with the levels of all the known sources of noise that contribute to it.

1.6.2.a Seismic and anthropogenic noise

Seismic noise, which arises from the sub-terranean motion of the Earth, is a major factor for any ground-based interferometer. Anthropogenic noise caused by human activity, especially from machinery, is also important in the audible band – both sources of noise primarily couple through the suspensions of the test masses and the beam splitter. To reduce this noise, the aLIGO-aVirgo detectors employ multi-stage suspension systems for the test masses and the beam splitter. In aLIGO , this takes the form of a quadruple pendulum system in which the test masses are the final stage. Each stage attenuates this noise by a factor of $1/f^2$, yielding an attenuation proportional to $1/f^8$. In combination with the active isolation system also used by aLIGO, this ensures that seismic noise only becomes an important source of noise at low frequencies, below ~ 10 Hz, as seen in Fig. 1.7. The reader is pointed to Ref. [66] for a detailed review of the techniques employed to acheive the necessary level of seismic isolation in aLIGO.

1.6.2.b Quantum Noise

Another important source of noise is the quantum noise stemming from the fundamentally quantum nature of the laser light. This can be divided into two types, photon shot noise and radiation pressure noise. Shot noise is associated with fluctuations in the laser power due to fluctuations in the number of photons in the laser beam. The amplitude spectrum of shot noise depends on frequency as follows:

$$S_{\rm shot}^{1/2}(f) = \frac{1}{8\mathcal{F}L}\left(\frac{4\pi\hbar\lambda_L c}{\eta P_{bs}}\right)^{1/2}\sqrt{1+\left(\frac{f}{f_p}\right)^2}, \tag{1.49}$$

where $P_{\rm bs}$ is the laser power at the beam splitter, λ_L is the laser wavelength, $L = 4$km is the arm length, $\mathcal{F}$ is the cavity finesse, f_p is the pole frequency of the cavity. During the third observing run f_p was 411 Hz for the LIGO-Hanford detector and 455 Hz for the LIGO-Livingston detector [4]. The laser power $P_{\rm bs} \approx 1500$ W at LIGO-Hanford and $P_{\rm bs} \approx 1800$ W at LIGO-Livingston. Both detectors use a Nd:YAG laser with $\lambda_L = 1064$ nm and employ a cavity Finesse $\mathcal{F} = 450$ in the advanced detector era [24].

The other type of quantum noise is the radiation pressure noise. This arises from the fluctuations in the radiation pressure on the test masses from the laser light which results

in a fluctuating force. The amplitude spectrum of this noise depends on frequency as,

$$S_{\rm rad}^{1/2}(f) = \frac{4\mathcal{F}}{\pi^2\sqrt{\pi}ML f^2}\left(\frac{\hbar P_{bs}}{\lambda_L c}\right)^{1/2}\frac{1}{\sqrt{1+\left(\frac{f}{f_p}\right)^2}}, \tag{1.50}$$

where M is the mass of the end mirrors. We thus see that while shot noise grows with frequency, the radiation pressure decreases. Furthermore, while shot noise decreases with increasing laser power, the radiation pressure noise increases. This leads to the idea of the standard quantum limit (SQL) which is the minimum bound on the total quantum noise at a given frequency; found by minimizing with respect to the power. This has deep ties with the Heisenberg's uncertainty limit, as is revealed by a more quantum-mechanical treatment of the noise [67, 68] in which shot noise and radiation pressure noise are related to each other as conjugate quantum-mechanical variables.

Fortunately, the SQL is not an absolute limit on the sensitivity of a GW detector. Rather, one type of quantum noise can be "squeezed" into another such that the total quantum noise is reduced below the SQL at a given frequency while still staying true to the limit imposed by the uncertainty principle [69]. Such "quantum squeezing" is achieved in practice by injecting squeezed vacuum states at the output port of an interferometer [70, 71]. Frequency independent quantum squeezing was used in both aLIGO and aVirgo detectors during O3 to reduce quantum shot noise without a corresponding increase in laser power [72, 5]. Shot noise is the primary limiting noise source at frequencies above 100 Hz as seen in Fig. 1.7 making it an apt target for quantum squeezing. This increased the average binary neutron star detection range of the Hanford and Livingston detectors by 12% and 14%, while the average range of the Virgo detector increased by $5-8\%$.

Figure 1.8 shows the broadband reduction in interferometric noise due to quantum squeezing that was achieved in the Livingston detector.

1.6.2.c Thermal noise

Thermal noise of various types are also relevant for interferometric detectors. This includes thermal noise in the test masses due to heating caused by the laser. A similar noise type is the thermal noise in the reflective coating on the mirrors and in the

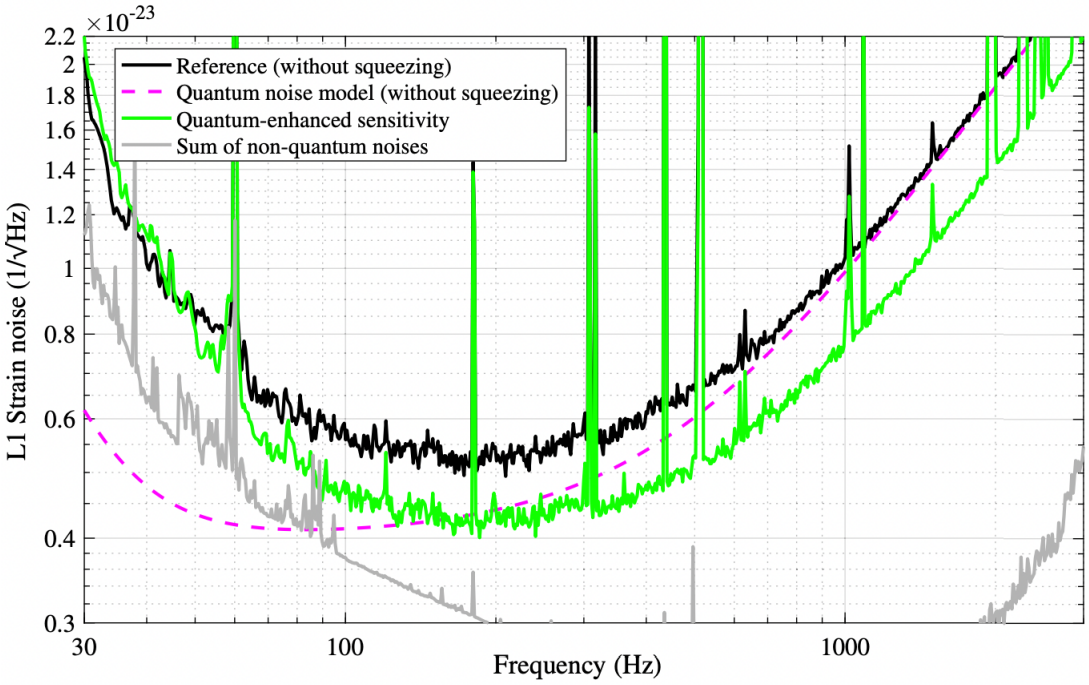


Figure 1.8: The reduction in interferometric noise levels at the Livingston detector during O3 due to quantum squeezing. Figure reproduced from Ref. [5] with permission.

suspension systems.

1.6.2.d Miscellaneous broadband noise sources

There is also scattered light noise due to the various reflections and scatterings, either on surfaces within the interferometer or on any residual gas within the arms. Furthermore, there are technical noises associated with the various control loops and servo systems used with the detector, and noises associated with laser frequency and amplitude fluctuations. There is also Newtonian noise, produced by local fluctuating gravitational influence on the test masses; for example by seismic and atmospherical motions. I point to Ref. [4] for a detailed overview of all these noise sources and their impact on the sensitivity and performance of the aLIGO detectors during O3.

1.6.2.e Narrowband sources of noise

In addition to broadband sources of noise, there are also narrow spectral lines caused by various types of resonances [6, 2]. Both the aLIGO detectors have loud instrumental lines at 60 Hz, and harmonics thereof due to the coupling from the North American AC power. There are also resonant modes of the test mass and beam splitter suspension, at $\sim$ 500 Hz and the harmonics thereof called Violin modes. These modes are especially strongly excited by seismic or anthropogenic noise which couple through the suspension systems, such as during earthquakes.

There are also various loud narrowband lines manually injected to assist with various control loops. Calibration lines are also added through the photon calibrator system [73] which provides the main calibration standard to convert the power output at the photodiode into dimensionless GW strain [74]. All identified narrowband lines are removed before any GW data analysis by a combination of Wiener filtering and band-rejection filtering.

1.6.2.f Glitches

A persistent noise and data quality issue is the occurrence of transient non-Gaussian artifacts called glitches in GW detectors. Glitches often introduce large amount of spectral power relative to a GW signal into the differential arm (DARM) Michelson channel. They are especially pernicious when they occur close by or overlap with a transient BBH signal. A loud glitch famously overlapped with the GW170817 BNS signal in the Livingston detector during the second observing run [25]. This delayed the dissemination of the full three-detector skymap, which was crucially important to discovering the electromagnetic counterpart, by several hours [27].

The source of some of these glitches is understood, and can be traced through environmental monitoring channels. Figure 1.9 shows spectrograms of some common glitch types. The rate of glitches was significantly higher in O3 compared to the earlier runs which necessitated gating algorithms or algorithms for modeling the glitches to ensure that they do not contaminate the GW searches [75, 76, 77].

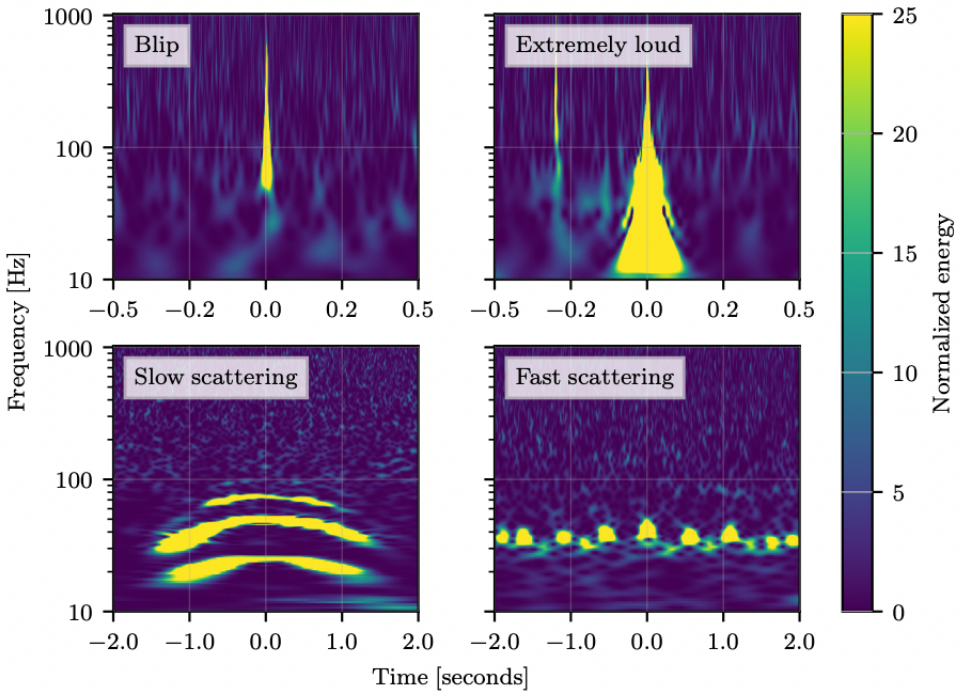


Figure 1.9: This figure, reproduced with permission from Ref. [6], shows Q-transform-based spectrograms of four commonly seen kinds of glitches. The sources of the blip and the extremely-loud glitches are not entirely understood. The slow and the fast scattering glitches on the other hand are caused by scattering light joining the main laser beams in the interferometer, and are excited by microseismic and anthropogenic activity respectively.

1.6.2.g Schumann Noise

A source of noise which is uniquely important to SGWB searches is due to Schumann resonances, which are magnetic resonances induced by lightning strikes in the resonant cavity created in the the Earth's atmosphere [7]. They can couple to GW detectors though cables, electromagnets and magnetic substances in the suspension systems. This creates a noise, which while weak can be correlated between widely-separated GW detectors. While they are unproblematic for most GW searches due to their weakness, because they show up as noise correlated between GW detectors, they are a significant noise source for SGWB searches. Figure 1.10 shows the Schumann resonant peaks

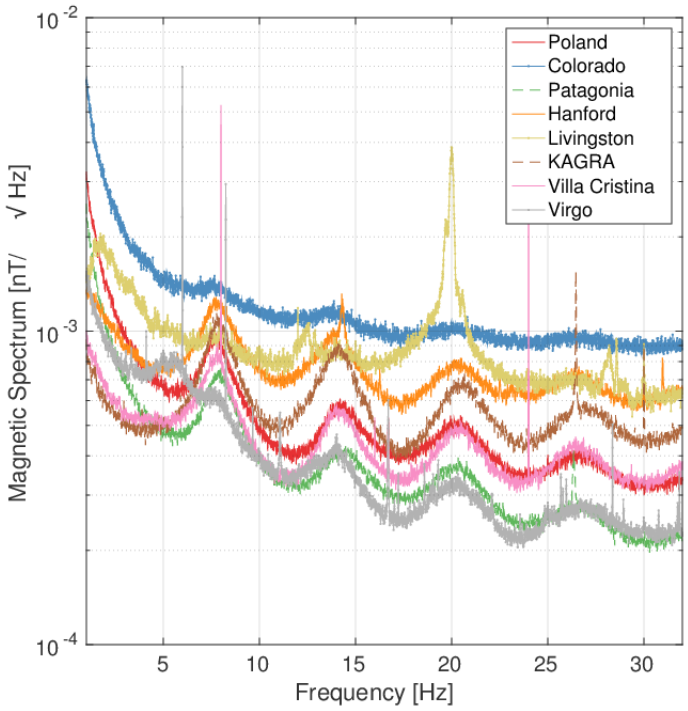


Figure 1.10: The figure shows Schumann resonances at $\sim 8, 14, 21, 27$ and 32 Hz by magnetometers located at various places on Earth. The figure is reproduced from Ref. [7] with permission.

showing up in widely separated interferometers across the Earth.

Various methods have been proposed for dealing with Schumann resonance-based correlated noise in SGWB searches, including Weiner Filtering [7] and Bayesian model selection [78]. The latter method, when used on data from the aLIGO-aVirgo interferometers, found no measurable sign of this correlated noise as of O3 [79] indicating that the searches are still dominated by instrumental noise. Nevertheless, correlated Schumann noise is expected to be a limiting noise source for SGWB searches as detector sensitivities improve.

1.7 Gravitational-wave Astrophysics and Cosmology

As of the writing of this dissertation, aLIGO and aVirgo have confidently detected 50 GW signals, 39 from the first half of the third observing run (henceforth O3a), 8 from the second run (henceforth O2) and 3 from the first observing run(henceforth O1) [80, 81]. Analysis of data from the second half of the third observing run is still ongoing.

All in all, so far 47 BBH mergers and 2 BNS mergers have been confidently detected. One of the BNS mergers, GW170817 [25, 27] was also the first astrophysical multi-messenger transient detected through GWs. The other BNS merger was detected in O3a and was effectively a single detector event since the Hanford detector was offline at that time and the signal was too weak to be seen in the Virgo detector [82]. The O3a run also saw the first GW detection of an intermediate mass black hole with GW190521 [83]. The mass of the remnant is estimated to be $\sim 140 M_{\odot}$ and lies at the lower end of the intermediate mass range which is thought to begin at $100 M_{\odot}$. A candidate electromagnetic counterpart was detected by the Zwicky Transient facility, potentially consistent with a BBH merger happening in the environment of an active galactic nucleus [84], although the statistical significance of this association is disputed [85, 86].

Following hierarchical population analyses of O1 + O2 + O3a data, the local BBH merger rate has been estimated to the $R_{\rm BBH}(z=0) = 23.9^{+14.9}_{-8.6}{\rm Gpc}^{-3}{\rm yr}^{-1}$ [87] at 90% credible intervals. With the detection of high mass events in O3, there is evidence for a high mass pile-up of black holes in the BBH mass distribution around $\sim 35 M_{\odot}$, corresponding to the effects of pulsation pair-instability supernova. As detector sensitivities increase and more detections happen, the population analysis will start to pin down on the features of binary distributions, exploring different formation channels which tie back in with the astrophysics of star formation.

The BNS merger rate is estimated to be $R_{\rm BNS} = 320^{+490}_{-240}{\rm Gpc}^{-3}{\rm yr}^{-1}$ [87], the larger error bars being indicative of the relative paucity of detected BNS mergers. The O2 event, GW170817, was one of the most well studied astrophysical events and it dominates our understanding of BNS mergers. Multi-messenger studies of GW170817 have provided evidence that r-processes happen in the kilonova following the merger (see Ref. [27] and the references therein) which is an astrophysical pathway for the production of the heavy elements of the periodic table. Comparison of the GW arrival time

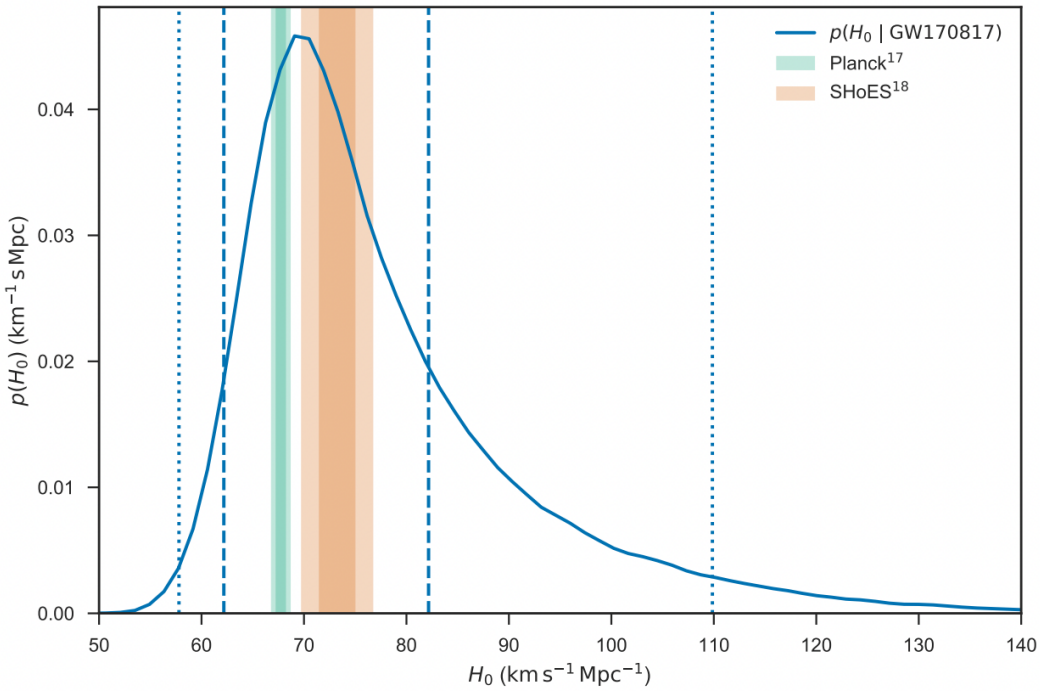


Figure 1.11: Posterior distribution on the Hubble constant obtained from the standard siren measurement of GW170817. The plot is reproduced from Ref. [8] with permission.

with that of the accompanying gamma ray burst, detected by Fermi Gamma-Ray Burst Monitor and INTEGRAL [88], has given the most precise limits on the deviation of the speed of GW propagation from that of light. By using waveforms incorporating matter effects, aLIGO-aVirgo have also been able to measure the tidal deformability of neutron star matter [89, 25].

There are also potential cosmological applications of detecting GWs. For example, locating the electromagnetic counterpart of the BNS merger GW170817 also allowed for a standard siren measurement of the Hubble constant at $H_0 = 70^{+12.0}_{-8.0}\mathrm{kms}^{-1}\mathrm{Mpc}^{-1}$ [8]. The Fig. 1.11 shows the posterior distribution of Hubble constant from this measurement. More multi-messenger BNS detections will allow for a more precise measurement of the Hubble constant, and potentially settle the current tension between local and early Universe measurements [90, 91, 92].

The measurement of merger rates from the GW signals detected thus far also allow

for an estimate of the level of the astrophysical SGWB formed from an incoherent superposition of unresolvable BBH and BNS mergers. The latest such calculation done by incorporating the merger rates from O3a estimate that $\Omega_{\rm BBH}(f = 25{\rm Hz}) = 5.0^{+1.7}_{-1.4} \times 10^{-10}$ and $\Omega_{\rm BNS}(f = 25{\rm Hz}) = 2.1^{+2.9}_{-1.6} \times 10^{-10}$ for the SGWBs from the unresolvable BBH mergers and BNS mergers respectively [79]. This places the total background at a factor of few from current upper limits, making the detection of this background potentially likely in the coming few years.

With rising sensitivities of ground-based detectors and the promise of space-based LISA detector, GW detections are allowing us to observe the Universe in unprecedented ways and open up new avenues for astrophysics and cosmology. The latter part of this dissertation, starting with Chapter 3, will deal with some applications of GWs towards those goals. While the problem considered in this dissertation are somewhat disparate, they are all united by the common theme of being scientific questions that will be uniquely accessible through GW channels.

Science may be described as the art of systematic over-simplification –– the art of discerning what we may with advantage omit.

Karl Popper

Chapter 2

Bayesian Inference

2.1 Introduction

Bayesian statistics has become an essential part of modern astrophysics and cosmology, playing a crucial connecting role between data and models. This chapter will provide a quick primer on Bayesian inference, focusing mainly on its applicability to GW physics. Bayesian statistics is interpreted through the lens of the Bayes theorem, which itself is an example of an application of conditional probabilities. Bayes theorem states that given data d, model parameters θ and background assumptions I_0, the posterior probability distribution $P(\theta|d, I_0)$ is the posterior probability distribution of θ given the data

$$P(\theta|d, I_0) = \frac{\mathcal{L}(d|\theta)\pi(\theta|I_0)}{\mathcal{Z}(d, I_0)}. \tag{2.1}$$

The goal of Bayesian inference often is the computation of the posterior to understand the inferences that can be made about the model parameters based on the data. The term $\mathcal{L}(d|\theta)$ in Eq. 2.1 is called the Likelihood and is the conditional distribution of the data given θ. It represents the probability of generating data d given parameters θ and thus captures the nature and statistics of the data generation process. The term $\pi(\theta|I_0)$ is the prior distribution, which can fold in any prior information or uncertainty regarding the values of the parameters θ. The prior and the posterior being probability

distributions of θ, need to be normalized as follows,

$$\begin{aligned} \int P(\theta|d, I_0)\, d\theta = 1, \\ \int \pi(\theta|I_0)\, d\theta = 1. \end{aligned} \tag{2.2}$$

The likelihood, on the other hand, is not necessarily normalized with respect to θ as it is better interpreted as the probability distribution of the data.

The denominator $\mathcal{Z}(d, I_0)$ in Eq. 2.1 is called the Bayesian evidence or the marginal likelihood. As the name suggests, it represents the probability of obtaining the data d once we marginalize or integrate over the uncertainty of θ. Therefore, it is purely a function of the data and the prior model used. While the evidence does not play an important role in calculating the posterior, it can be essential in comparing hypotheses or in hierarchical inference. It is generally calculated as the integral of the likelihood weighted by the prior,

$$\mathcal{Z}(d, I_0) = \int \mathcal{L}(d|\theta)\pi(\theta|I_0)\, d\theta. \tag{2.3}$$

This can be seen from normalizing Eq. 2.1 with respect to θ. From here on we will implicitly assume the background information I_0 and suppress it unless necessary.

2.2 Model Selection

Bayesian model selection gives the odds [1] between two or more hypotheses based on the data, giving us a metric to test hypotheses and see how much the data prefers one model over the other. The odds ratio $\mathcal{O}_{\alpha\beta}$ between two models α and β is the ratio of the posterior [2] probabilities of the models i.e

$$\mathcal{O}_{\alpha\beta} = \frac{P(\alpha|d)}{P(\beta|d)} = \frac{P(d|\alpha)\,\pi_\alpha}{P(d|\beta)\,\pi_\beta} = \frac{\mathcal{Z}_\alpha(d)}{\mathcal{Z}_\beta(d)}\frac{\pi_\alpha}{\pi_\beta}, \tag{2.4}$$

where $\pi_\alpha : \pi_\beta$ gives the prior odds for the two hypothesis. In many cases, the prior odds is taken to be one, in which case the odds ratio is just the ratio of evidences, also

[1] In the same sense as betting odds.

[2] Note that the posteriors being described here are for the entire models, marginalized over the uncertainties in model parameters. They are not the same posteriors described in Eq. 2.1.

$\mathcal{B}_{\alpha\beta}$	$2\ln\mathcal{B}_{\alpha\beta}$	Odds for α over β
< 1	< 0	Model β preferred over α
$1-3$	$0-2$	Barely worth a mention
$3-20$	$2-6$	Positive odds
$20-150$	$6-10$	Strong odds
> 150	> 10	Very strong odds

Table 2.1: Interpretation of Bayes factors. Adapted from Ref. [1]

known as the Bayes factor $\mathcal{B}_{\alpha\beta}$,

$$\mathcal{O}_{\alpha\beta} = \mathcal{B}_{\alpha\beta} \equiv \frac{\mathcal{Z}_\alpha(d)}{\mathcal{Z}_\beta(d)}. \tag{2.5}$$

Table 2.1 provides reference values to interpret the odds given by the Bayes factor. However, one should note that the table is only to be interpreted as a rule of thumb. Bayes factors represent odds ratios between competing hypotheses, and while it might favor one hypothesis over the other, the notion of "ruling out a hypothesis" should not be applied here.

There are several other ways to compare competing hypotheses besides using Bayes factors. Some of the other common methods include the Akaike information criterion, the likelihood ratio, Bayesian information criterion, each with its strengths and weaknesses. One of the advantages of using Bayes factors (or odds ratios) is that they perfectly fit into the Bayesian formalisms without any additional assumption. Another significant advantage is that the notion of "Occam's razor" is built into them by design. In other words, if two competing hypotheses, α and β, fit the data equally well, the Bayes factor will always favor the model with a fewer number of parameters. This can be easily seen in the specific example worked out in the Laplace approximation in Ref. [93]. However, calculating the evidence by marginalizing the likelihood can be a computational challenge beyond some simple models. Section 2.4.2 will touch on statistical nested samplers, which present one way to estimate the evidence for complicated models.

2.3 Likelihood function for gravitational-wave analysis

To utilize Bayesian statistics, we need to model the statistics of the data generation process – both of the instrumental noise and the signal if applicable – through the likelihood function. This section will describe the basic likelihood function used for GW astrophysics and derive some of the more specific likelihoods that will be useful for this dissertation. It should be noted at the offset that the problems discussed in this chapter represents only a narrow slice of the range of applicability of Bayesian ideas in GW and multi-messenger astrophysics. For an excellent overview of the full breadth of their application, the reader is pointed to [94].

It is usually convenient to work in the Fourier domain when working with data from laser interferometers, wherein we assume that the instrumental noise of the interferometer, $\tilde{n}(f)$ is generated through a colored stationary Gaussian process with zero mean and with the one-sided power spectral density (PSD) $S(f)$. Since $\tilde{n}(f)$ is a complex number that implies that both the real and the imaginary parts of the noise are independent and identically distributed $(i.i.d)$ as Gaussian with zero mean and variance $S(f)/2$,

$$\begin{aligned} P(\tilde{n}_{\mathbb{R}}(f)) &= \sqrt{\frac{2}{\pi T_{\rm seg} S(f)}}\, \exp\left(-\frac{2\,\tilde{n}^2_{\mathbb{R}}(f)}{T_{\rm seg} S(f)}\right), \\ P(\tilde{n}_{\mathbb{I}}(f)) &= \sqrt{\frac{2}{\pi T_{\rm seg} S(f)}}\, \exp\left(-\frac{2\,\tilde{n}^2_{\mathbb{I}}(f)}{T_{\rm seg} S(f)}\right), \end{aligned} \tag{2.6}$$

where the discrete Fourier transform was assumed to be calculated using segment duration $T_{\rm seg}$. The extra factor of $2/T_{seg}$ is proportionality factor between raw discrete Fourier transforms and the one-sided PSD. Combining them together, we can write a probability distribution function for $\tilde{n}(f)$, sometimes called the Whittle likelihood [93, 11]

$$P(\tilde{n}(f)) = \frac{2}{\pi T_{\rm seg} S(f)}\, \exp\left(-\frac{2|\tilde{n}(f)|^2}{T_{\rm seg} S(f)}\right), \tag{2.7}$$

Therefore, in the presence of a GW signal $\tilde{h}(f,\theta)$ in the data $\tilde{d}(f)$, we can write a likelihood function based on the residual,

$$\mathcal{L}(\tilde{d}(f)\,|\,\theta) = \frac{2}{\pi T_{\rm seg} S(f)} \exp\left(-\frac{2|\tilde{d}(f) - \tilde{h}(f,\theta)|^2}{T_{\rm seg} S(f)}\right). \tag{2.8}$$

Note that $\tilde{h}(f,\theta)$ is not the metric perturbation, but is rather the strain in the detector given by Eq. 1.33. It, therefore, contains the action of the antenna patterns upon the GWs. When working with multiple frequencies, one can simply use the orthogonality assumption of the Whittle likelihood to get,

$$\mathcal{L}(d\,|\,\theta) = \prod_f \frac{2}{\pi T_{\rm seg} S(f)} \exp\left(-\frac{2|\tilde{d}(f) - \tilde{h}(f,\theta)|^2}{T_{\rm seg} S(f)}\right), \tag{2.9}$$

or in terms of the log-likelihood as,

$$\log \mathcal{L}(d\,|\,\theta) = -\sum_f \frac{2|\tilde{d}(f) - \tilde{h}(f,\theta)|^2}{T_{\rm seg} S(f)} - \sum_f \log\left(\pi T_{\rm seg} S(f)/2\right). \tag{2.10}$$

The summation here is only over positive frequencies since we use the one-sided PSD $S(f)$ and $\tilde{n}(-f) = \tilde{n}^*(f)$ for real time-domain data, where the asterisk implies complex conjugate.

In the presence of n GW detectors or multiple channels, we promote the likelihood to a multivariate complex Normal and the spectral density to a $n \times n$ covariance matrix $\tilde{C}(f)$. The diagonal terms of this matrix contain the PSD of the individual channels, while the off-diagonal terms capture the cross-spectral density (CSD), which is a measure of the covariance. Equation. 2.10 is then written as,

$$\begin{aligned}\log \mathcal{L}(d\,|\,\theta) = -\frac{2}{T_{\rm seg}} \sum_f \left(\tilde{d}^*(f) - \tilde{h}^*(f,\theta)\right) \tilde{C}^{-1}(f) \left(\tilde{d}(f) - \tilde{h}(f,\theta)\right) \\ - \sum_f \log\left(\pi T_{\rm seg}^n |\tilde{C}(f)|/2\right).\end{aligned} \tag{2.11}$$

The term $|\tilde{C}(f)|$ is the determinant of the covariance matrix (as a function of frequency). In the case of the current generation ground-based detectors, which are spatially separated, one usually assumes, after appropriate data-quality checks, that the noise in the detectors is independent of each other and that all the non-diagonal terms in the

covariance matrix are zero. In which case, one can simply write the log-likelihood as a summation over detectors or channels I in addition to frequencies

$$\log \mathcal{L}(d \,|\, \theta) = -\frac{2}{T_{\rm seg}} \sum_I \sum_f \frac{|\tilde{d}_I(f) - \tilde{h}_I(f,\theta)|^2}{S_I(f)} - \sum_I \sum_f \log\left(\pi T_{\rm seg} S_I(f)/2\right). \quad (2.12)$$

This expression forms the basis for much of the CBC parameter estimation algorithms [95, 96, 97] and the analyses done on the GW catalogs thus far [80, 81].

2.3.1 Stochastic likelihood

The likelihood defined in Eq. 2.8 and beyond assumes that the signal model is well known, depending on parameters θ. While this is true for CBC and continuous-wave signals, it is not so for other GW categories as indicated by Fig. 1.4. In particular, an SGWB is only defined by its statistical properties, and it is not possible to have a signal template by definition. An SGWB is generally assumed to be stationary, unpolarized, and colored Gaussian, defined by a one-sided PSD $S_{\rm gw}(f)$, which is the only measurable quantity. We can thus write the probability distribution of $\tilde{h}'(f)$, the Fourier domain amplitude of the SGWB, treating it as a statistical realization based on $S_{\rm gw}(f)$ [3],

$$p(h'(f)) = \frac{2}{\pi T_{\rm seg} S_{\rm gw}(f)} \exp\left[-\frac{2|\tilde{h}'(f)|^2}{T_{\rm seg} S_{\rm gw}(f)}\right]. \quad (2.13)$$

where $\tilde{h}'$ is meant to differentiate it from the strain in the detector $\tilde{h}(f)$ which depends also on the antenna patterns (see Eq. 1.33). Since the only measurable quantity is $S_{gw}(f)$ we need the likelihood $\mathcal{L}(\tilde{d}(f) \,|\, S_{\rm gw}(f))$, which can be derived through using conditional probabilities as [98],

$$\mathcal{L}(\tilde{d}(f) \,|\, S_{\rm gw}(f)) = \int dh'\, P(\tilde{d} \,|\, h) p(h') \quad (2.14)$$

[3] This expression suppresses the polarization state of the SGWB for purposes of clarity. But explicitly writing the polarization will just give us a product of two Gaussians and the same math carries over

Using Eq. 2.8 here (and suppressing frequency dependence within the integral for clarity), we get,

$$\begin{aligned}\mathcal{L}(\tilde{d}(f)\,|\,S_{\rm gw}(f)) &= \int_{-\infty}^{\infty} dh \frac{2}{\pi T_{\rm seg} S} \exp\left[-\frac{2|\tilde{d}-\tilde{h}|^2}{T_{\rm seg}S}\right] \frac{2}{\pi T_{\rm seg} S_{\rm gw}} \exp\left[-\frac{2|\tilde{h}'|^2}{T_{\rm seg}S_{\rm gw}}\right], \\ &= \frac{2}{\pi T_{\rm seg}\left(\mathcal{R}(f)S_{\rm gw}(f)+S(f)\right)} \exp\left[-\frac{2|\tilde{d}(f)|^2}{T_{\rm seg}\left(\mathcal{R}(f)S_{\rm gw}(f)+S(f)\right)}\right],\end{aligned} \tag{2.15}$$

which is just a complex Gaussian with a PSD given by $\mathcal{R}(f)S_{\rm gw}(f) + S(f)$. This is consistent with the fact that the sum of two random variates is itself a random variable with a variance that is the sum of the original two. The term $\mathcal{R}(f)$ is the response function of the detector to GW power. For an isotropic, unpolarized SGWB $\mathcal{R}(f)$ given by,

$$\mathcal{R}(f) = \sum_A \int d^2\hat{\mathbf{n}}\, |F^A(f,\hat{\mathbf{n}})|^2. \tag{2.16}$$

One can again generalize Eq. 2.15 to multiple frequency bins and n detectors or channels as,

$$\begin{aligned}\log \mathcal{L}(d\,|\,S_{\rm gw}) = &-\frac{2}{T_{\rm seg}} \sum_f \left[\tilde{d}^*(f)\left(\tilde{C}_{\rm gw}(f)+\tilde{C}(f)\right)^{-1}\tilde{d}(f)\right] \\ &- \sum_f \log\left(\pi T_{\rm seg}^n |\tilde{C}_{\rm gw}(f)+\tilde{C}(f)|/2\right),\end{aligned} \tag{2.17}$$

where $\tilde{C}_{\rm gw}(f)$ is an $n \times n$ covariance metric which, in addition to the strength and the spectral shape of the SGWB, also captures details of how it couples with the detectors.

2.3.2 Marginalization over PSD Uncertainty

The likelihood in Eq. 2.3 depends on the noise statistics of the GW detector, in particular its PSD, $S(f)$. However, the noise PSD is not apriori known in the presence of a signal since any estimate will also contain power from a GW. Also complicating any such estimation is the fact that the instrumental noise is not entirely stationary and can

fluctuate over long timescales. One thus needs a way of either measuring or modeling the noise PSD for Bayesian inference (and in most frequentist algorithms too). Some algorithms, such as Bayeswave, take the latter route [99] and fit the noise PSD in the Bayesian process using a series of noise splines as models. A more common approach, used in frequentist and Bayesian transient GW searches alike, is to estimate the noise PSD using neighboring data segments [4] under the assumptions that they contain no signal and that the noise is stationary over the data stretch being used. A common approach is then to average the PSD of some N neighboring data segments.

$$S_{\rm avg}(f) = \frac{1}{N}\sum_{i}^{N} S_i(f). \tag{2.18}$$

Even when the assumption of stationarity is valid, any such estimate will come with an uncertainty, which, if unaccounted for, can bias our inference. The noise in the frequency domain $\tilde{n}(f)$ is complex, with both the real and imaginary parts independently drawn from colored Gaussian noise with variance $S(f)/2$. Therefore the sum,

$$\frac{2}{S(f)}\sum_{i}^{N} S_i(f) = \frac{2N}{S(f)} S_{\rm avg}(f), \tag{2.19}$$

follows a χ^2 distribution with $2N$ degrees of freedom. We can thus account for the uncertainty in the PSD by marginalizing over the actual PSD $S(f)$ as follows (and once again suppressing the frequency dependence within the integral for the sake of clarity),

$$\mathcal{L}(\tilde{d}(f)\,|\,\theta) = \int dS\, P(\tilde{d}\,|\,\theta, S)\, p(S|S_{\rm avg}), \tag{2.20}$$

where $p(S|S_{\rm avg})$ is given by the χ^2 distribution,

$$p(S|S_{\rm avg}) = \chi^2_{2N}\left(\frac{2NS_{\rm avg}(f)}{S(f)}\right) \tag{2.21}$$

In general if a random variable $X \sim N(0,\sigma^2)$ and $\nu Y^2/\sigma^2 \sim \chi^2_\nu$, where Y is an estimator for σ, then the random variable $t = X/Y$ will form a Student-t distribution with ν degrees of freedom [100]. Thus we get the PSD marginalized likelihood in Eq. 2.20

[4] The length of the data segment in the time domain is equal to the duration used for the discrete Fourier transform.

to be:

$$\mathcal{L}(\tilde{d}(f)\,|\,\theta) = \frac{4\,\Gamma\left(1+\frac{\nu}{2}\right)}{\pi T_{\rm seg}\nu S(f)\,\Gamma(\nu/2)}\left[1+\frac{4}{T_{\rm seg}}\frac{|\tilde{d}(f)-\tilde{h}(f)|^2}{\nu\,S(f)}\right]^{-\left(1+\frac{\nu}{2}\right)}, \tag{2.22}$$

with the natural choice of $\nu = 2N$. Note that since we start with the complex Gaussian distribution Eq. 2.8, the exponent is not $-(\nu+1)/2$ as would be expected from a standard Student-t distribution but rather is $-\left(1+\frac{\nu}{2}\right)$.

2.4 Statistical sampling methods

The central problem of Bayesian inference is mapping the posterior in a multidimensional space of parameters θ. An analytical mapping, however, is impossible in but a few rare cases. Calculating the Bayesian evidence through the integral in Eq. 2.3 can get even more complicated. For very low dimensional spaces, one can solve the problem numerically by calculating the posterior probability density on points along a grid in the parameter space. However this quickly becomes computationally untenable as the dimensionality of the parameter space increases. For example, suppose that one is working in a four-dimensional parameter space and the grid needs a 1000 points on each dimension. Then, the total number of points at which the likelihood has to be evaluated rises to 10^{12}. And usually most astrophysical and cosmological problems have many more parameters to estimate than four [101]. This is sometimes called the "curse of dimensionality".

To overcome this challenge, one can use statistical sampling or Monte Carlo techniques. Roughly speaking, these methods attempt to make a number of fair draws from a (posterior) probability distribution we want to analyze, without having to calculate an analytical form for it. The number of points in a region will be proportional to the probability density in that region. Thus the probability distribution can be reconstructed by histograms or by kernel density estimation.

Suppose the sampler draws samples θ_i from the probability distribution $p(\theta)$. Let $g(\theta)$ be some other function with a finite expectation value. Then, the strong law of large numbers guarantees that the mean over the draws θ_i converges to the expectation

value, i.e.,

$$\lim_{n\to\infty} \frac{1}{n} \sum_{i=1}^{n} g(\theta_i) \to \int d\theta g(\theta) p(\theta) \tag{2.23}$$

With a sufficiently large but finite number of draws, one can approximate this as,

$$\frac{1}{n} \sum_{i=1}^{n} g(\theta_i) \approx \int d\theta g(\theta) p(\theta). \tag{2.24}$$

This expression captures the fundamental utility of statistical samplers. Any quantity that depends on θ and its uncertainty can be estimated using a finite number of samples. Moreover, by definition the tails of the probability distribution $p(\theta)$ contribute much less to such an estimate compared to the peak of the distribution. Therefore the number of fair draws needed for an estimate is usually much lower than what would be needed for a uniform grid on the parameter space.

2.4.1 Markov chain Monte Carlo samplers

The most popularly used statistical samplers belong to two categories, one of them being Markov chain Monte Carlo (MCMC) samplers. MCMC samplers work by moving a random walker through the parameter space of a probability distribution $p(\theta)$. The walker forms a Markov chain, i.e., its transition probability from the current point θ_0 to a point θ_1 depends only on θ_0 and θ_1.

The first implementation of an MCMC sampler was through the Metropolis algorithm, later extended by Hastings and now known as the Metropolis-Hastings (MH) algorithm [102, 103]. In the MH sampler, the transition probability is related to a proposal distribution $q(\theta'|\theta_0)$ from which a proposal point θ' is drawn [5]. The proposal should be symmetric i.e $q(\theta'|\theta_0) = q(\theta_0|\theta')$. We also draw a random variable $\mathcal{U}$ from the uniform distribution between 0 and 1. If

$$\frac{p(\theta')}{p(\theta_0)} > \mathcal{U}, \tag{2.25}$$

[5]The proposal function can be any arbitrary continuous function whose analytical form is known in principle. However, in practice, for the sampler to properly explore the parameter space will require an appropriate choice of the transition function suited for the kind of probability distribution at hand. Running an MCMC sampler is partly an art.

the new proposed point is accepted and $\theta_1 = \theta'$; otherwise $\theta_1 = \theta_0$. This process is repeated several times to obtain a *chain* which asymptotically approaches the target probability distribution. There are several variations of MCMC sampling, which are offshoots of the Metropolis-Hastings algorithms such as Gibbs sampling [104], Hamiltonian MCMC [105].

Most MCMC samplers suffer from the problem that the first few samples will strongly depend on the initial condition of the Markov Chain and often require a burn-in phase. The parts of the chain generated during this phase are rejected. Moreover, since the sampler operates as a Markov chain, subsequent draws of the chain are not *i.i.d* but are instead highly correlated with each other. To mollify this, one often thins an MCMC chain with relation to autocorrelation metrics. Finally, one can never be sure that an MCMC sampler has fully explored all the modes of a probability distribution in a finite time. This can be fixed somewhat by running multiple independent walkers, initialized independently to run in parallel.

Some of the popular MCMC samplers used in astrophysics include emcee [106], pymc3 [107] and stan [108]. MCMC sampling algorithms are also implemented in the LALInference package that is used by LIGO and Virgo for analyzing GW data [95].

2.4.2 Nested sampling

The other type of statistical sampling technique that is popularly used is Nested sampling. Unlike MCMC samplers, the main goal of nested samplers is not just to explore the posterior but calculate the Bayesian evidence through Eq. 2.3. The approach to nested sampling was first described by Skilling [109].

Nested samplers generally walk through the prior distribution as opposed to MCMC samplers which directly target the posterior. A predefined number of *live points*, n, are first drawn from the prior distribution, and the likelihoods are calculated at those points. The live point with the lowest likelihood, say θ_0, is discarded, and another point θ_1 is drawn subject to the constrain that $\mathcal{L}(\theta_1) > \mathcal{L}(\theta_0)$ [6]. This process is iterated, each time discarding the point with the lowest likelihood. This builds up a nested set of samples around regions of high likelihood. One can statistically compute the fractional evidence contained within each of the nested steps, allowing us to estimate the evidence integral

[6] Most nested sampling algorithms use some variety of MCMC step to sample the new point

in a probabilistic (Monte Carlo) manner.

While calculating the evidence is the primary function of the nested sampler, we also get equal-weighted posterior draws as a happy by-product. One can do this by weighing the set of points - both discarded and otherwise - by the posterior values at those points.

Some of the popularly used nested samplers include Dynesty [110], MultiNest [111, 112] and Ultranest [113, 114]. Nested sampling algorithms are also implemented in the LALInference package that is used by LIGO and Virgo for analyzing GW data [95].

2.5 Hierarchical Bayesian modeling

The notion of Bayesian inference discussed so far dealt with a single data generating process, either transient or continuous. We can, however, take this one level further to attempt to infer the population distribution of a set of such data generating processes – which we will call events. If one can model such a distribution, the individual events are assumed to be *i.i.d* drawn from it [7]. Hierarchical Bayesian inference has played a crucial role in using the GW detections by aLIGO-aVirgo to understand the astrophysical distribution of their parameters and learn about their formation channels [115, 87].

If we have an event with data d, one extends Bayes theorem to accommodate population properties as,

$$P(\theta, \Theta|d) = \frac{\mathcal{L}(d|\theta)\pi(\theta|\Theta)}{\mathcal{Z}(d)}\pi(\Theta), \tag{2.26}$$

where $\pi(\Theta)$ is called the hyper prior and describes the distribution of hyperparameters Θ. The hyper prior models the population distribution from which individual events are drawn. The prior on the intrinsic parameters for one event, $\pi(\theta|\Theta)$, is now a conditional distribution on the hyperparameters.

The posterior, $P(\theta, \Theta|d)$, is now a distribution over not just the intrinsic parameters θ, but also over the hyperparameters Θ. The posterior over the hyperparameters alone

[7] One can, of course, go more levels up the hierarchical ladder and talk about populations of populations, etc.

can be obtained by marginalizing over the intrinsic parameters,

$$P(\Theta|d) = \int d\theta P(\theta, \Theta|d) = \frac{\mathcal{L}(d|\Theta)\pi(\Theta)}{\mathcal{Z}(d)}, \tag{2.27}$$

where we have defined the marginal likelihood,

$$\mathcal{L}(d|\Theta) = \int d\theta \mathcal{L}(d|\theta)\pi(\theta|\Theta) \tag{2.28}$$

In the non-hierarchical case, we generally refer to such an integral as the evidence, but in this case, it is more aptly referred to as a likelihood since it depends on the hyperparameters Θ. With multiple *i.i.d* events from the same population, labeled by $t = \{1, 2, ..., N\}$, we obtain posteriors for the hyperparameters by simply taking the product over the events,

$$P(\Theta|d) \propto \pi(\Theta) \prod_{i=1}^{N} \mathcal{L}(d_i|\Theta) \tag{2.29}$$

On a final note, the discussion here has omitted selection effects as they are not relevant to the problems considered within this dissertation. However, selection effects play an important role in many applications of hierarchical Bayesian inference to GW data and astrophysics at large. The reader is directed to [116] for more details about selection effects, but also hierarchical inference in general in the GW context.

2.5.1 Recycling

The Eqs. 2.26 and 2.27 describe how the posteriors distributions of the hyperparameters Θ can be calculated. Notice, however, that the (regular) likelihood function $\mathcal{L}(d|\theta)$, usually the most computationally expensive piece to calculate, does not depend on hyperparameters at all. When using a statistical sampler, this property lets us, in some cases, to apply Bayesian inference on two separate levels. This technique called recycling or sample recycling allows us to massively reduce the computational expense of running a statistical sampler in the hierarchical case.

Suppose one has equal-weighted fair draws on the intrinsic parameters θ using a prior $\pi(\theta|\Theta_0)$ with fixed hyper-parameter values Θ_0, and now desires to use a different prior model $\pi(\theta|\Phi)$ to infer hyperparameters Φ. The marginal likelihood using the new

prior is defined by using Eq. 2.28:

$$\mathcal{L}(d|\Phi) = \int d\theta \mathcal{L}(d|\theta)\pi(\theta|\Phi) \tag{2.30}$$

However, one can connect the marginal likelihood through Bayes theorem to the posterior probability density at the same θ but with some other hyperparameters Θ_0 [8]:

$$\mathcal{L}(d|\theta) = P(\theta|d, \Theta_0)\frac{\mathcal{L}(d|\Theta_0)}{\pi(\theta|\Theta_0)}, \tag{2.31}$$

where we have used marginal likelihood in place of evidence, sticking with the previous convention. Then we get,

$$\mathcal{L}(d|\Phi) = \mathcal{L}(d|\Theta_0)\int d\theta P(\theta|d, \Theta_0)\frac{\pi(\theta|\Phi)}{\pi(\theta|\Theta_0)}. \tag{2.32}$$

Since we have assumed that we already have fair draws from $P(\theta|d, \Theta_0)$, we use Eq. 2.24 to get

$$\mathcal{L}(d|\Phi) = \frac{\mathcal{L}(d|\Theta_0)}{n}\sum_{i=1}^{n}\frac{\pi(\theta_i|\Phi)}{\pi(\theta_i|\Theta_0)}, \tag{2.33}$$

where the summation is over the previously sampled posterior samples. Recycling thus allows us to calculate the marginal likelihood with different priors without having to redo the statistical sampling, effectively decoupling the problem of inference of hyperparameters from a computational perspective. A couple of caveats apply. The original prior $\pi(\theta_i|\Theta_0)$ must allow for the at least the same range in θ as $\pi(\theta_i|\Phi)$. Moreover, because the drawing of the posterior samples itself depends on the prior used, $\pi(\theta_i|\Theta_0)$ also needs to be generally broad and not sharply peaked at any point in the θ space.

[8] This is because the marginal likelihood plays the same role as the evidence for the intrinsic parameters if we fix the hyperparameters.

Chapter 3

Search for long-lived post-merger signals after GW170817

This chapter can be divided into two parts. The first part, in Sec. 3.1 and 3.2 gives a brief introduction to BNS post-merger remnants and briefly describes results of the various post-merger searches that were conducted by the LIGO-Virgo collaboration in the wake of the BNS merger GW170817. The second part, from Sec. 3.3 onwards, takes a deep dive into the Stochastic Transient Analysis Multidetector Pipeline (STAMP) search for long-lived post-merger remnants from GW170817. Four other pipelines also took part in this study; however, the chapter mainly focuses on the STAMP search, which was the main contribution of this author. The results were published in a paper in the Astrophysical Journal [3] by the LIGO-Virgo collaboration.

3.1 Introduction

The BNS merger GW170817 [25] was an exceptional astrophysical event. It remains, to date, the only GW event to be also conclusively observed through electromagnetic channels [27] (although there are other claims [84]). Section 1.7 described some of the astrophysical and cosmological implications of the multimessenger observations of GW170817.

The BNS merger was also an extremely close event in astronomical terms, at just

40 Mpc away from us. While the GW localization was crucial to pinpointing the electromagnetic counterpart, the localization volume obtained through GW data was fairly broad [25, 27]. This will generally be true of all CBC detections with the current ground-based detectors; the posterior sky and distance localizations for such signals will be at least a few square degrees and a few Mpc, respectively. However, the identification of the transient AT 2017gfo in the galaxy NGC 4993 with the electromagnetic counterpart of GW170817 gives us a precise 3D localization of the event [117, 118, 119]. The close distance and the precise localization make GW170817 an ideal testbed to look for post-merger GW signals and develop algorithms that can target these kinds of signals.

The fate of the post-merger remnant of a BNS merger depends not only on its mass but also on the factors such as the equation of state, the progenitor mass ratio, and other details [120, 9, 121, 122, 123]. The different paths of remnant evolution possible from a BNS merger are schematically shown in Fig 3.1. The heaviest remnants will undergo a prompt collapse into a black hole. Somewhat lighter remnants with masses below the collapse threshold will form a hypermassive neutron star, temporarily supported against gravitational collapse by differential rotation, where the outer layers rotate slower than the inner layers. Remnants that are even lighter can be supported by uniform rigid body rotation and are called supramassive neutron stars. While the former is expected to collapse within a second or so, the latter can remain stable for thousands of seconds after the merger before eventual collapse [1]. The lightest remnants can eventually end up as stable neutron stars.

We expect some post-merger GW radiation in all of these scenarios, either from the ring down of the prompt-collapse black hole remnant or from the dynamics of the neutron star remnants as they slow down. The ringdown GWs from a small remnant black hole are predominantly expected to be at frequencies of several kHz [122]; too high to be detectable by ground-based detectors. On the other hand, GWs from the various neutron star remnants, while still at high frequencies, are detectable by ground-based interferometers in principle. Henceforth, the name post-merger will only refer to the material remnant phases unless otherwise specified.

The post-merger GW waveform is expected to be complicated and difficult to model

[1] Note that the hypermassive and supramassive neutron stars are more accurately described as phases in the evolution rather than final states. As their angular momentum is radiated away, they will settle as black holes or neutron stars.

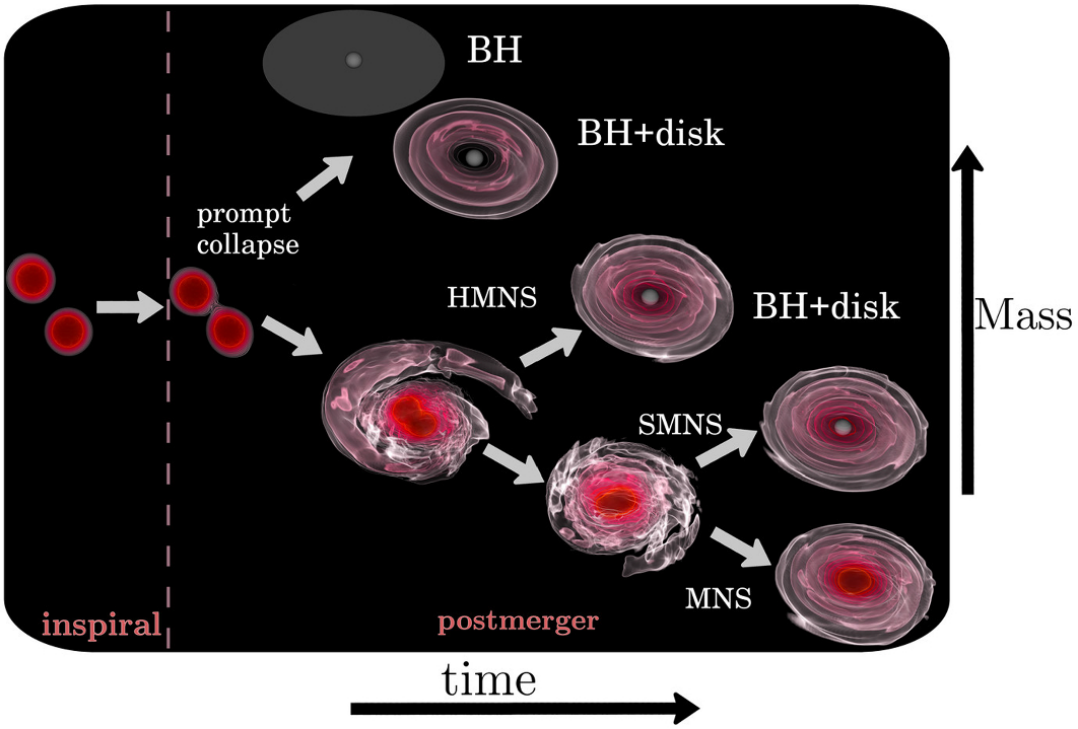


Figure 3.1: A schematic of the different types of merger remnants possible. Reproduced with permission from Ref. [9].

accurately, depending on details of relativistic magnetohydrodynamics, thermal effects, and nuclear physics (e.g., Refs. [124, 125, 126, 127, 128, 129]). Nevertheless, GWs from a post-merger remnant present a rare opportunity to study the physics of the densest and more energetic forms of nuclear matter at strong gravity. In particular, we could constrain the EoS of matter at conditions inaccessible even within regular neutron stars, making these remnants attractive laboratories for otherwise inaccessible physics. Since accurate modeling is difficult, we rely on unmodeled algorithms that make minimal assumptions about the waveform to search for such signals.

3.2 LIGO-Virgo post-merger searches

Following GW170817, there was a concerted effort to search for post-merger GWs and to help develop search strategies for the future. In order to cover all categories of signals,

multiple pipelines were employed, each focusing on GWs of specific range of durations. There was about 8.5 days of data after GW170817 until the end of the O2 run. While no post-merger signals were found in this data by any of the pipelines, they were able to set upper limits and show that about an order of magnitude increase in strain sensitivities over O2 was needed at high frequencies ($\gtrsim 1$ kHz) to start being sensitive to post-merger signals.

On the short end, the Bayeswave [99] and coherent WaveBurst [130] pipelines were used to search for GW signals from a potential hypermassive neutron star remnant. The signal durations considered for these searches were about $\mathcal{O}(1)$ s after the coalescence of GW170817 [10, 131]. While both pipelines are unmodeled and operate in the wavelet basis, they employ different philosophical approaches to detection. The coherent WaveBurst pipeline is frequentist in nature while Bayeswave is a Bayesian algorithm that uses a trans-dimensional reversible-jump MCMC algorithm to simultaneously fit GW signals and the instrumental noise. Figure 3.2 shows the full frequency-dependent limits set by Bayeswave using data from both the LIGO detectors and the GEO 600 detector [10]. The coherent WaveBurst analysis meanwhile set optimal upper limits of $h_{\rm rss}^{50\%} = 2.1 \times 10^{-22}$ Hz$^{-1/2}$ [2] between $1 - 4$ kHz [131].

The search for GWs from long-lived remnants was done in two separate iterations. The coherent waveBurst and the Stochastic Transient Analysis Multidetector Pipelines (STAMP) [132] were used to search for post-merger GWs from a supramassive remnant with a duration between $\sim 10 - 1000$ s. Both are unmodeled searches that do not make specific assumptions about the nature of the waveform. They set optimal limits on post-merger radiation, from a magnetar spin-down model at $h_{\rm rss}^{50\%} = 8.4 \times 10^{-22}$ Hz$^{-1/2}$ and at $h_{\rm rss}^{50\%} = 5.9 \times 10^{-22}$ Hz$^{-1/2}$ for a bar-mode model [131].

Then, there were searches searches that targeted long-lived post-merger signals that lasted more than a thousand seconds. Four different pipelines were used here, STAMP, a hidden Markov-model (HMM) based search pipeline [133, 134, 135] and two search pipelines based on the Hough transform [136]; the Generalized FrequencyHough [137] and Adaptive Transient Hough pipeline [138]. These four pipelines have complementary strengths and made different modeling and analysis assumptions but ultimately yielded

[2] $h_{\rm rss}$ is the root mean square strain defined as $h_{\rm rss} = \sqrt{2\int_{f_{\rm min}}^{f_{\rm max}} df \left(|\tilde{h}_+(f)|^2 + |\tilde{h}_\times(f)|^2\right)}$.

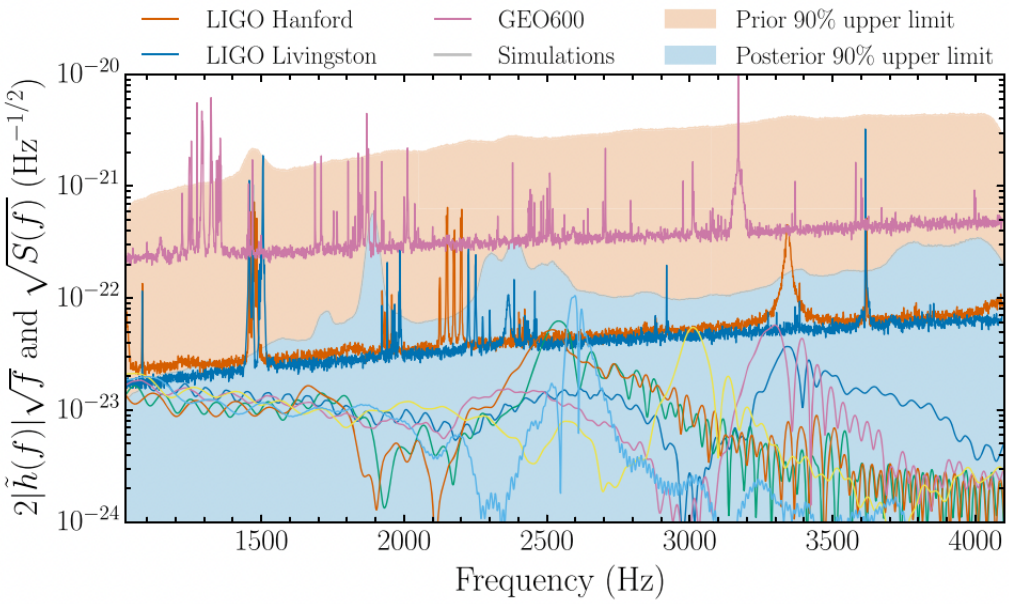


Figure 3.2: Upper limits for the Bayeswave post-merger analysis following GW170817, as a function of frequency. The noise ASDs of the three detectors are plotted as solid lines while the orange and cyan regions correspond to the 90% limits on the amplitude prior and the amplitude posterior densities. Results for several numerical simulations are shown for comparison. Figure reused from Ref. [10] with permission.

results that are consistent with each other. The rest of this chapter will mainly focus on the application of the STAMP pipeline towards the search for long-lived post-merger remnants.

3.3 STAMP

The STAMP pipeline [132] works with time-frequency maps (TF-maps) constructed by cross-correlating data from multiple GW detectors. Cross-correlation is a powerful technique that can help extract weak signals from noise. It is widely used for stochastic searches and searches for continuous-wave signals in LIGO-Virgo searches (for e.g. [139, 79, 140]), especially in the frequency domain. STAMP extends the method to search for GW transients by using it over TF-maps of data, which is particularly useful as an unmodeled search pipeline.

Consider the data in detectors I and J; they are linear combinations of the instrumental noise in the respective detectors and the strain from a potential GW signal

$$\begin{aligned} d_I(t) =& n_I(t) + h_I(t), \\ d_J(t) =& n_J(t) + h_J(t). \end{aligned} \tag{3.1}$$

The GW signal between the two detectors is correlated by nature, while we can usually assume that the instrumental noise is not. Thus, correlating the two data streams gives us,

$$\begin{aligned} \langle d_I(t) d_J(t) \rangle &= \langle n_I(t) n_J(t) \rangle + \langle h_I(t) n_J(t) \rangle + \langle n_I(t) h_J(t) \rangle + \langle h_I(t) h_J(t) \rangle, \\ &\approx \langle h_I(t) h_J(t) \rangle, \end{aligned} \tag{3.2}$$

where the terms $\langle h_I(t) n_J(t) \rangle$ and $\langle n_I(t) h_J(t) \rangle$ drop out because a GW signal and instrumental noise are uncorrelated by definition. Using the linear nature of the Fourier transform, the same cross-correlation property can be defined in the frequency domain, too i.e.

$$\langle \tilde{d}_I^*(t) \tilde{d}_J(t) \rangle \approx \langle \tilde{h}_I^*(f) \tilde{h}_J(f) \rangle. \tag{3.3}$$

The plane-wave GW signal from a source in the direction $\hat{\mathbf{n}}$ in the TT gauge is given by Eq 1.12, reproduced below

$$h_{ij}(t, \mathbf{x}) = \sum_{A=+,\times} \int_{-\infty}^{\infty} df\, \tilde{h}_A(f)\, e_{ij}^A(\hat{\mathbf{n}}) \exp\left\{-2\pi i f (t - \hat{\mathbf{n}} \cdot \mathbf{x}/c)\right\}. \tag{3.4}$$

Following Eq. 1.35, the signal in the detector I located at $\mathbf{x}_I$ with antenna pattern functions $F_I^A(t, \hat{\mathbf{n}})$ [3] is,

$$h_I(t) = \sum_{A=+,\times} \int_{-\infty}^{\infty} df\, \tilde{h}_A(f)\, F_I^A(t, \hat{\mathbf{n}}) \exp\left\{-2\pi i f (t - \hat{\mathbf{n}} \cdot \mathbf{x}_I/c)\right\}. \tag{3.5}$$

[3] in the large-wavelength limit appropriate for LIGO and Virgo

Let us take a finite-duration Fourier transform of $h_I(t)$ over some duration T to get,

$$\tilde{h}_I(t;f) = \sum_A \tilde{h}_A(f)\, F_I^A(t,\hat{\mathbf{n}})\, e^{2\pi i f \hat{\mathbf{n}}\cdot\mathbf{x}_I/c}. \tag{3.6}$$

This unit of Fourier-domain data, with a size of T seconds and (usually) $1/T$ Hz, is called an TF-map pixel. If we correlate these finite-duration pixels of data between two detectors I and J, then following Eq. 3.3 we get,

$$\langle \tilde{d}_I^*(t;f)\tilde{d}_J(t;f)\rangle \approx \langle \tilde{h}_I^*(t;f)\tilde{h}_J(t;f)\rangle = \sum_{A,A'} \langle \tilde{h}_A^*(f)\tilde{h}_{A'}(f)\rangle F_I^A F_J^{A'} e^{2\pi i f \hat{\mathbf{n}}\cdot\Delta\mathbf{x}_{JI}/c}, \tag{3.7}$$

where $\Delta\mathbf{x}_{JI} = \mathbf{x}_J - \mathbf{x}_I$ is the spatial separation between the detectors.

It is usually a good assumption to take the GW to be unpolarized when using an unmodeled pipeline. In which case the correlation can be written as,

$$\langle \tilde{d}_I^*(t;f)\tilde{d}_J(t;f)\rangle \approx \frac{1}{2} H(t;f) \sum_A F_I^A\, F_J^A\, e^{2\pi i f \hat{\mathbf{n}}\cdot\Delta\mathbf{x}_{JI}/c}, \tag{3.8}$$

where $H(t;f)$ is the instantaneous (over time scales of duration T) power of the GW transient. Since the power is a real number, the complex phase of the correlation will be entirely given by $2\pi i f \hat{\mathbf{n}}\cdot\Delta\mathbf{x}_{IJ}/c$. This motivates the use of an estimator $\hat{Y}$ for the TF-map pixel given by i.e.

$$\hat{Y}(t;f;\hat{\mathbf{n}}) = 2\,\Re\left[\tilde{Q}(t;f,\hat{\mathbf{n}})\, \tilde{d}_I^*(t;f)\, \tilde{d}_I(t;f)\right], \tag{3.9}$$

where $\tilde{Q}(t;f,\hat{\mathbf{n}})$ is a filter function whose phase nullifies the GW phase from the cross-correlation. It is mathematically given by,

$$\tilde{Q}(t;f,\hat{\mathbf{n}}) = \frac{2\, e^{-2\pi i f \hat{\mathbf{n}}\cdot\Delta\mathbf{x}_{JI}/c}}{\sum_A F_I^A\, F_J^A}. \tag{3.10}$$

Thus, the filter function ensures that the detector pair is "pointed" in the right direction during the search and that the estimator $\hat{Y}$ is appropriately normalized. If the assumed direction is close to the correct direction, the complex phase cancels out, and the Y will roughly be equal on average to the GW power – a positive quantity – as can be seen

from combining Eqs. 3.8 and 3.9,

$$\langle \hat{Y}(t; f; \hat{\mathbf{n}}) \rangle \approx H(t; f). \tag{3.11}$$

If the direction is completely incorrect, the estimator $\hat{Y}$ will become negative for a GW signal be.

The variance of the $\hat{Y}$ can be calculated in the weak signal limit to be [132],

$$\sigma^2_{\hat{Y}}(t; f, \hat{\mathbf{n}}) \approx \frac{1}{2} |\tilde{Q}(t; f, \hat{\mathbf{n}})|^2 \, P_I(t; f) P_J(t; f), \tag{3.12}$$

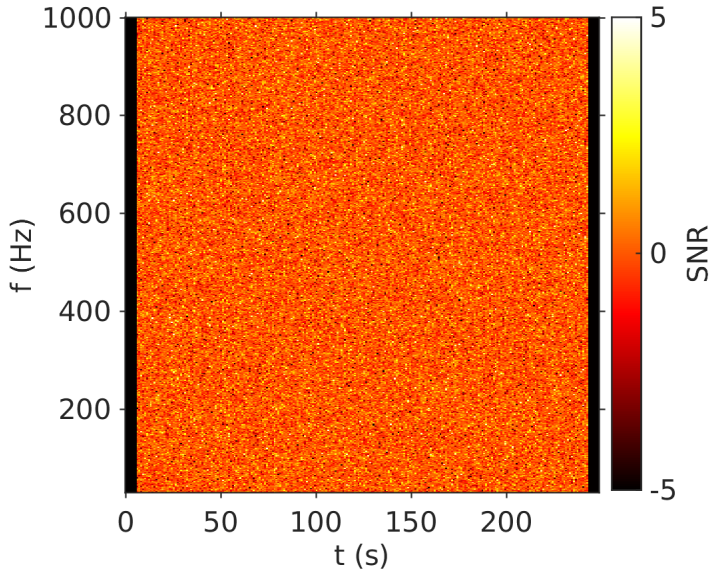
where $P_I(t; f)$ is the PSD of the detector noise. One thus defines a signal-to-noise ratio (SNR) of the TF-map pixel as,

$$\rho(t; f, \hat{\mathbf{n}}) = \frac{\hat{Y}(t; f; \hat{\mathbf{n}})}{\sigma_Y(t; f, \hat{\mathbf{n}})} = \Re \left[2\sqrt{2} \, \frac{\tilde{Q}(t; f, \hat{\mathbf{n}}) \, \tilde{d}^*_I(t; f) \, \tilde{d}_I(t; f)}{|\tilde{Q}(t; f, \hat{\mathbf{n}})| \sqrt{P_I(t; f) P_J(t; f)}} \right]. \tag{3.13}$$

The *pixel SNR* $\rho(t; f, \hat{\mathbf{n}})$ is the basic unit of analysis in the STAMP pipeline. The statistical distribution of the pixel SNR in the absence of a signal is well approximated as a unit Normal with zero mean, but the actual distribution follows a modified Bessel function of the second kind [141]. The reader is directed to Ref. [132, 141] for more details about the STAMP algorithm.

3.3.1 SNR TF-maps

The pixel SNR $\rho(t; f, \hat{\mathbf{n}})$ defined in Eq. 3.13 allows us to create SNR TF-maps with several desirable properties. The distribution of pixels in these SNR TF-maps closely follow a unit Normal with zero mean in the absence of a signal or any instrumental noise transients. The pixel SNRs are real, unlike FFTs, and yet unlike the PSD spectrograms, contain information about the directionality of the signal due to the filter function. A signal will show up as a cluster of pixels with positive SNR in the TF-map. Figure 3.3 shows two SNR TF-maps; the top plot is made purely from simulated noise from LIGO Hanford (henceforth H1) and LIGO Livingston (henceforth L1) detectors. The bottom subplot has a loud simulated signal added to the noise, which appears as a narrowband track on the SNR TF-maps. Thus, STAMP reduces the problem of detecting transient GWs to a problem of pattern recognition.



(a)

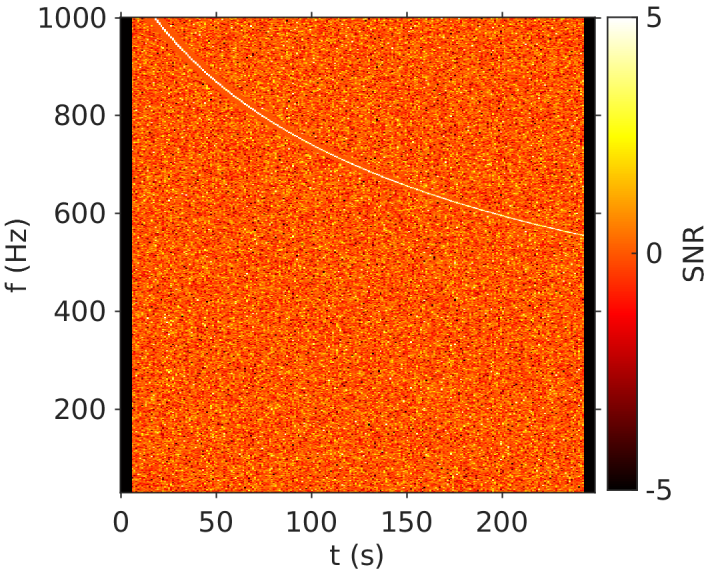


Figure 3.3: The plots show two SNR TF-maps generated from simulated H1 and L1 noise with O2 detector sensitivity. The map on the top contains only detector noise, while an artificial, loud magnetar spin-down signal has been added to the bottom map. The signal shows up as a bright track on the map.

3.4 Pattern recognition algorithms

Several pattern-recognition algorithms have been proposed and developed for identifying GW clusters in SNR TF-maps. The first STAMP paper proposed an algorithm based on the Radon transform, designed to identify line-like features [132].

An important class of clustering algorithms that were developed for this problem are seeded algorithms [142, 143]. These algorithms start by identifying loud pixels and building clusters using the pixels surrounding these bright pixels. Seeded clustering algorithms are, therefore, completely unmodeled by design. However, since they start with bright pixels, they can be insensitive towards GWs that are not locally loud but still stand out against the noise. For instance, a long-lived GW signal will not create individual pixels that are very bright since the power at any given instant will be low. However, its statistical significance is tied to the fact that it lasts very long and will stay above the background.

3.4.1 Seedless clustering

Seedless clustering algorithms circumvent these issues by not needing any initial seeds at all. Instead, in seedless algorithms tracks drawn from a bank of templates are overlaid on the TF-map to approximate and pick out the morphology of the signal. These templates could be very generic, or they can be specific templates that incorporate physical information about the signals. For example, if one were trying to detect CBC signals through seedless clustering, the natural bank of templates will be those generated from CBC waveform models [144].

A more generic template model can robustly fit more kinds of signal tracks and can act as an unmodeled clustering algorithm. In particular, the Stochtrack version of the seedless clustering algorithm uses quadratic Bézier curves on the TF-map as a template back [145, 146]. These curves provide good sensitivity to most conceivable monotonically evolving narrowband GW signal models.

The way these curves are drawn is as follows. Three pixels in a SNR TF-map $N_i = (t_i, f_i), i \in \{0, 1, 2\}$ are randomly chosen with the condition that they be monotonically placed in frequency. Then a quadratic Bézier curve connecting them can be

parameterized by,

$$\Big(t(\xi), f(\xi)\Big) = (1-\xi)^2 N_0 + 2(1-\xi)\xi N_1 + \xi^2 N_2. \tag{3.14}$$

Each different selection of three points yields a different Bézier curve. Generally, $\mathcal{O}(10^6)$ templates are used for a single TF-map, and the best curve – as chosen by a detection statistic – is considered the GW trigger for an TF-map. Section 3.4.2 will discuss construction and usage of detection statistics with STAMP. The STAMP pipeline with the Stochtrack algorithm can be used over GW data as a directed search [145], targeting sources whose location is well known, or as an all sky-search [146]. It was used for various transient searches during both the initial and advanced detector phase [147, 148, 149, 131, 150]. Stochtrack is also the clustering algorithm used by the STAMP pipeline in the search for a potential long-lived post-merger remnant following GW170817 [3], as described by the rest of this chapter.

One should note that while the Bézier curve-based Stochtrack algorithm is usually more sensitive [145], it is by design less robust than seeded clustering. For example, it can be less sensitive to signal with a more broadband spectrum, like GWs from a core-collapse supernova. It also struggles with signals that are poorly approximated by the parabola on the TF-map. This was demonstrated by a comparison of its performance to a Hidden Markov Model-based algorithm in Ref. [151].

3.4.2 Detection statistic

Whatever the clustering algorithm of choice, its output is a set of pixels in the SNR TF-map, denoted by Γ. One needs a detection statistic to assess the statistical significance of the cluster and the sensitivity of the algorithm. It is natural to adopt the weighted sum of SNRs of the individual pixels of the cluster as a statistic

$$\rho_\Gamma = \sum_{i\in\Gamma} w_i \rho_i. \tag{3.15}$$

ρ_Γ is often called the trigger or cluster SNR. For the GW170817 long-lived post-merger search, all the weights are chosen to be equal and depend on the length of the cluster such that

$$\rho_\Gamma = \frac{1}{N^{3/4}} \sum_{i \in \Gamma} \rho_i, \tag{3.16}$$

where N is the number of pixels in the cluster. This inverse weight of $N^{3/4}$ strikes a balance that ensures that longer signals are not too preferably picked over shorter ones just from noise SNR fluctuations, while also ensuring that longer signals are not suppressed. The detection statistic is calculated for all the $\mathcal{O}(10^6)$ Stochtrack tracks returned from an TF-map, and the track with the highest ρ_Γ is chosen as the trigger for the TF-map.

The significance of this trigger is determined by comparing its ρ_Γ against the background distribution, which in turn is calculated by running the pipeline on data where one of the detectors is artificially time-shifted with respect to the other. The light (and GW) travel time between the LIGO detectors is of the order of 10 ms. Hence, any time shift larger than that will destroy correlations between the GW signal in one detector to another. The triggers generated from such time-shifted data will be generated by noise fluctuations alone [4]. Sufficient numbers of such noise triggers will let us understand the background noise statistics against which the *on-source triggers* which might contain a real GW signal are weighed for calculating statistical significance.

3.5 STAMP analysis for a long-lived remnant

The rest of this chapter will describe the specifics of the STAMP search by the LIGO-Virgo collaboration for long-lived post-merger remnants following GW170817 [3].

3.5.1 Configuration

The STAMP pipeline was run with the Stochtrack algorithm, with a focus on signals with durations anywhere from $\mathcal{O}(1000\text{ s}) - \mathcal{O}(100{,}000\text{ s})$ long. To be sensitive to such long-lived but weak signals, the pipeline was run in a novel configuration with spectrograms that are 15,000 s long. The search was directed at sky-position of the electromagnetic counterpart of GW170817, i.e, (ra, decl) = (13.1634 Hrs, $-23.3185°$).

[4] after sufficient data-quality cuts.

The search itself was performed in two versions, a low frequency version using TF-maps between 30 to 2000 Hz, and a high frequency version between 2000 to 4000 Hz. The rationale for searching up to such high frequencies is that the compact object left over from the merger will be rapidly spinning due to the conservation of angular momentum. Since the predominant GW emission mode is expected to be quadrupolar, with frequencies twice that of the rotational frequency, extending the search to the kilohertz band is appropriate. Nevertheless, a search at such high frequencies is challenging from a data quality perspective. The instrumental noise level steadily rises as one goes to higher frequencies, reducing the detector sensitivity. Data calibration and characterization also become challenging at higher frequencies. For these reasons, the search was restricted to frequencies less than 4 kHz.

The low and high frequency versions have different pixel sizes. The low frequency TF-maps used fast Fourier transforms (FFT) of 100 s. The natural frequency resolution of such an FFT is 0.01 Hz, but the TF-map was coarse-grained to a frequency resolution of 1 Hz. This reduces the number of pixels in an TF-map [5] and is a necessary compromise for computational purposes to ensure that $\mathcal{O}(10^6)$ Bézier tracks can provide adequate coverage over an TF-map that is 15,000 seconds long. Coarse graining, however, hurts the sensitivity to signals that do not show reasonable frequency evolution over the duration of an TF-map.

The high frequency TF-maps are constructed from FFTs that use 50 s of data, then coarse-grained to 1 Hz. The choice of smaller pixels is driven by the necessity to minimize the loss in SNR of a potential high-frequency signal due to Earth's rotation. The filter function $\tilde{Q}$ that "points" the detector pair to a certain direction of the sky, defined in Eq. 3.10, is constant for any one pixel. However, the Earth's rotation slightly displaces the target sky position, and this displacement within a pixel made with large FFT durations can be enough to lower the SNR of the pixel. This limits the pixel duration in a frequency-dependent way. The FFT duration bound for a pixel for frequency f is given by [152]

$$T \leq \frac{\cos^{-1}(1-L)}{2\pi f(t_0)(\pi/24\,\mathrm{hr})}, \tag{3.17}$$

[5] By a factor of 100 of course.

where L is the fraction of the SNR that is lost due to Earth's rotation and $t_0 \approx 10\,\text{ms}$ is the light travel time between detectors. If we require L to be less than 10%, this limits the the segment duration to be $T \leq 50\text{s}$ at 4000 Hz, which is thus the chosen FFT duration for the high frequency search.

3.5.2 Data-quality

The on-source part of the search consists of about 8.5 days of O2 data after GW170817 until the end of the O2 run, from 17 to 25th August 2017. The background run that generated noise triggers used data from 24th July 2017 to just before the merger on 17th August (see Sec. 3.5.3). The STAMP search only used data from the H1 and L1 aLIGO detectors, and this was true of the other three searches in the paper as well. This is because the sensitivity of Virgo was significantly lower during O2 and there was also no science data from it before August 1, 2017 [3].

However, the time-series data is not one continuous chunk. There are intervals when a detector would not be taking science data, either because of maintenance or instrumental tests or significant instrumental noise levels – for example, due to an earthquake – or simply due to a loss of lock in the detector. In addition, data with significant broadband noise contamination were also removed based on pre-defined bad GPS time cuts. The duty cycle of the H1 and L1 detectors for the on-source part of the search was 83% and 85%, respectively, yielding a coincidence duty cycle of 75%.

Narrowband and broadband noise of instrumental or environmental origin also contaminate the data from each detector. Many of these narrowband lines are loud and would present an extended background to the searches if not removed. Some lines form a repeating comb occurring at fixed frequency intervals. Many lines, such as calibration lines, in addition to more broadband noise like beam jitter, were removed by the process of Wiener filtering [153]. Some lines were removed within the STAMP pipeline by notching them out, i.e., by setting the values of the pixels corresponding to these frequencies in the coarse-grained TF-maps to NaNs [6]. Table 3.1 lists the lines notched out in the STAMP pipeline and potential source for them, identified mainly by comparing to Ref. [2].

The nature and cause of some of these lines in the 2000 to 4000 Hz has not been

[6] This implies that the 1 Hz pixels are notched out.

perfectly identified. While some can be identified as harmonics of known lines at lower frequencies, many are unrelated. The data in this frequency band is generally not utilized by most searches, and instrumental checks are harder to do than at lower frequencies. So a conservative approach was adopted for this band by notching out loud frequency lines and bands that show up in time-shifted data. It is important to note that since the data used for the background step is from before the coalescence we can be certain that any the notched lines are not from post-merger gravitational radiation, even without certain knowledge of their source. This (and all components of the background analysis) were done in a blinded way without revealing the results of the on-source runs to reduce human bias.

As discussed in Sec. 1.6.2.f, GW detectors often witness broadband transient non-Gaussian artifacts called glitches. These glitches are usually localized to less than a second in time, but the degree of power contained in them can significantly affect any detection algorithm. STAMP uses an algorithm called glitch cut to identify and remove glitches before using clustering algorithms [154]. The central idea behind the glitch cut algorithm is the fact that it is statistically unlikely for both detectors to have coincident glitches. In other words, a glitch can be detected as excess power in one detector compared to the other, and the algorithm constructs a statistic called glitch SNR for this purpose. The reader is directed to Ref. [154] for more details about the construct of the glitch SNR statistic.

While glitch cut works well for shorter duration maps that were used for previous STAMP searches, coincident glitches become more likely with longer TF-maps employed here. An additional narrowband cut was developed to reject these rarer coincident glitches. As the name suggests, the narrowband cut expects that any GW signals are narrowband in nature, a good assumption for the long-lived signals that are being targeted here.

The cut works as follows. It first identifies time segments that have an absolute cross-power SNR sum greater than a threshold ρ_{Tr},

$$\sum_{f} |\rho(t; f)| > \rho_{\mathrm{Tr}}. \tag{3.18}$$

The segment is identified as a glitch if the adjacent segments have absolute SNR sums

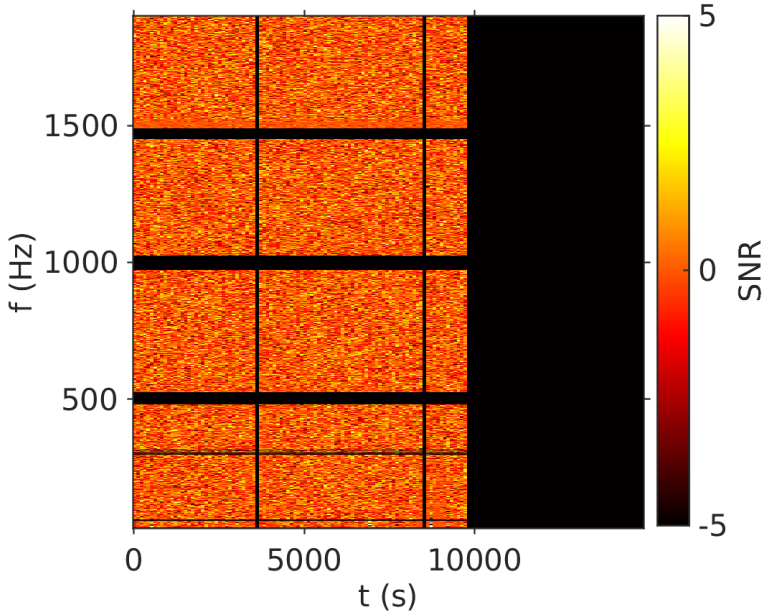


Figure 3.4: This 15,000 s long TF-map made from data from August 17th, 2017 depicts the effects of the data quality cut. The black vertical lines correspond to two glitches that have been removed. The horizontal black lines are the various narrowband features that have been notched out. Particularly prominent are the notched bands that correspond to the harmonics of the test-mass violin modes at ∼500, 1000, and 1500 Hz (see Tab. 3.1). The absence of data for the last ∼ 5000 s of the TF-map is due to a loss of lock at the L1 detector.

below this threshold. The comparison step is necessary to prevent loud signals or simulated signals from triggering the cut. The threshold value, after tests with simulated signals, was set at $\rho_{\mathrm{Tr}} = 700$. After glitches are identified, either by glitch cut or by narrowband cut, the time segments containing them are completely removed by changing the SNR values of all the pixels in those segments to NaNs.

Fig 3.5.2 shows an example of the effects of frequency notches, glitch rejection, and absence of data due to detector lock loss in a single SNR TF-map.

Line Frequency in Hz	Potential source	Detector of origin
30 - 33	Magnetic contamination and/or pulsar lines	H1, L1
58 - 62	60 Hz power line	H1 and L1
299, 300 - 303	Beam-splitter violin mode	H1
306 - 308, 315	Beam-splitter violin mode	L1
482 - 525	Test-mass violin modes	H1 and L1
972 - 1025	Test-mass violin modes (1st Harmonic)	H1 and L1
1450 - 1490	Test-mass violin modes (2nd Harmonic)	H1 and L1
1922 - 2036	Test-mass violin modes (3rd Harmonic)	H1 and L1
2107, 2108	unknown	unknown
2124 - 2126	unknown	unknown
2149, 2152	unknown	unknown
2224 - 2226	unknown	unknown
2249 - 2251	unknown	unknown
2274 - 2276	unknown	unknown
2298 - 2301	unknown	unknown
2336 - 2372	unknown	unknown
2373 - 2375	unknown	unknown
2380 - 2415	unknown	unknown
2419 - 2599	Test-mass violin modes (5th Harmonic)	H1 and L1
2821 - 2887	Test-mass violin modes (5th Harmonic)	H1 and L1
2918 - 2922	Test-mass violin modes (5th Harmonic)	H1 and L1
2985 - 3060	Test-mass violin modes (5th Harmonic)	H1 and L1
3220 - 3336	unknown	unknown
3384 - 3386,3397	Test-mass violin modes (6th Harmonic)	H1 and L1
3557 - 3620	Test-mass violin modes (6th Harmonic)	H1 and L1
3781 - 3793	unknown	unknown
3912 - 3923	Test-mass violin modes (7th Harmonic)	H1 and L1
3940 - 3943	Test-mass violin modes (7th Harmonic)	H1 and L1

Table 3.1: List of all the lines that were notched from the STAMP analysis in 1 Hz bands. The sources of the lines below 2 kHz are listed following Ref. [2]. The source of lines above 2000 Hz is sparse as most searches do not use data in this frequency band, and instrumental efforts too are more focused at frequencies below 2 kHz. To ensure that instrumental artifacts do not contaminate the search, notches were more liberally used in this band, using the background as a metric. As discussed above, the fact that these frequency notches were developed using data from before the merger guarantees their safety.

3.5.3 Significance of on-source triggers

Using the data quality cuts described above, H1 and L1 data from July 24 - August 17, 2017, was analyzed to determine the background for the search. In the background analysis, the data stream in one detector was shifted in time with respect to the other in intervals of 15,000 s [7]. At each offset, the data was divided into 15,000 s long TF-maps, overlapping with half of their neighbors in time. The triggers generated by the STAMP pipeline from the time-shifted TF-maps statistically represent the output we should expect if we run the search over interferometric noise alone. Since the analysis was conducted in separate low frequency and high-frequency modes, there are, in actuality, two independent sets of background triggers.

The rationale for using the data from before the merger for the background step was twofold. First, there was only about ~ 6.3 days worth of usable coincident data after the merger, which might not be enough to assess the significance of the on-source triggers confidently. Second, unlikely though it might be, we want to avoid potential correlations between possible post-merger GWs in the time-shifted data. Using pre-merger data for the background also allows us to be much more liberal with frequency notches, as described earlier, because no post-merger signal from GW170817 can exist before coalescence. The noise properties of the pre-merger data were inspected and found to be reasonably similar to the data after the merger when analyzed over scales of $\mathcal{O}(10{,}000)$ s.

Following the background runs, the on-source run was conducted using data coincident between H1 and L1 from after the merger in the same configuration as the background run. The data were again divided into 15,000 s long TF-maps with a 50% overlap with the adjacent maps. The same data quality cuts were applied as in the background runs, and on-source triggers were generated. Similar to the background run, there are separate on-source triggers for the high frequency and the low frequency components of the search.

Figure 3.5 shows the loudest trigger SNR ρ_Γ found in the on-source TF-maps, plotted on top of the cumulative ρ_Γ curve from the background. These triggers have p-values of 0.81 and 0.80, respectively, for the low and high-frequency searches as measured by

[7] This is after an initial offset of 15,000 s from coincidence. Coincident data is, of course, only used in the on source step to actually search for GW signals.

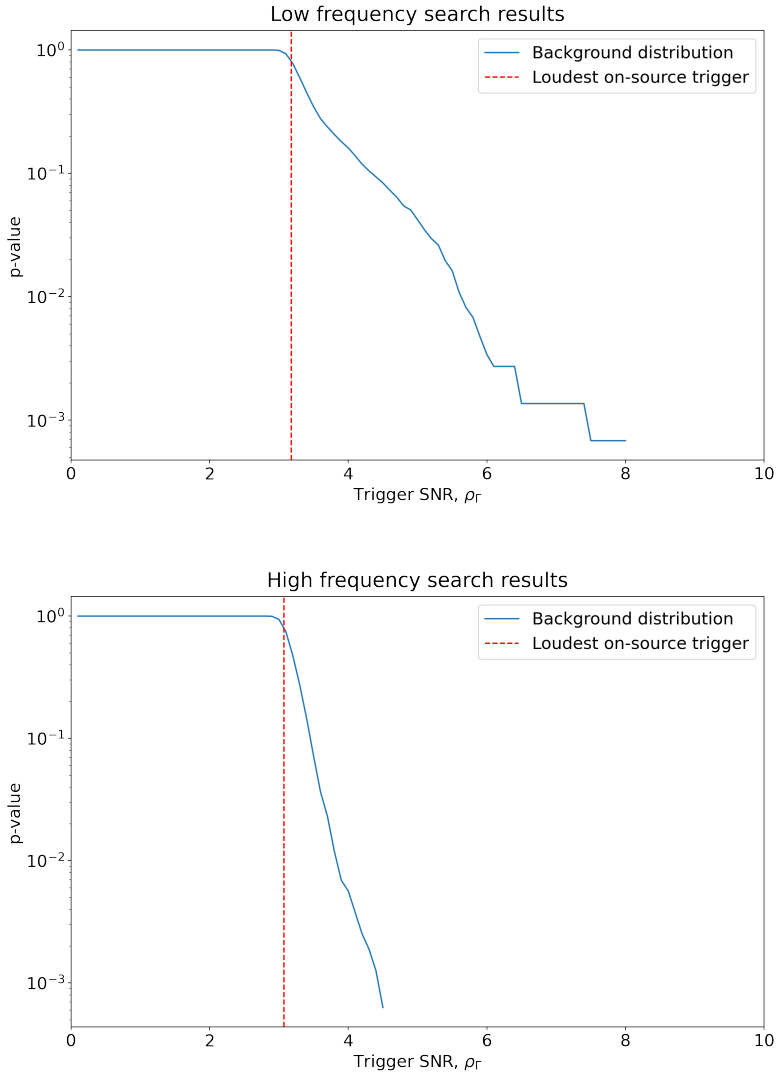


Figure 3.5: The plots show the ρ_Γ of loudest on-source trigger (also called zero lag) as a red dashed line; for the 30 to 2000 Hz band on the top, and the 2000 to 4000 frequency band on the bottom. The blue curve is the cumulative fractional distribution of the background triggers that are greater than or equal to a particular value. The p-value, i.e., the probability that the detector noise alone could have generated the on-source triggers, can be read off as the y-coordinate of the point where the dashed line intersects the curve. These triggers have p-values of 0.81 and 0.80, respectively. The data corresponding to this figure is from Ref. [3], reused here with permission.

the backgrounds, indicating that they are not statistically significant and driving us to conclude that the STAMP search did not find any long-lived post-merger GW signal following GW170817. This is broadly consistent with the results of the other three searches of the paper, none of which found a significant post-merger GW trigger [3].

3.6 Limits

With no evidence of long-lived post-merger GWs, we set limits on the strength of the possible signal that could go undetected. While STAMP is an unmodeled pipeline, a signal model is needed in order to set limits. For the long-lived post-merger search, STAMP and the three other pipelines adopted a spin-down model [155].

3.6.1 Spin-down magnetar model waveform

The model is based on the torque decay equation, which is often used in pulsar evolution studies,

$$\frac{d\omega}{dt} = -k\omega^n, \tag{3.19}$$

where ω is the angular velocity of the remnant, n is the braking index, and k is a constant of proportionality. The braking index is determined by the physics of the mechanism that dissipates energy and angular momentum. If the spin down is dominated by magnetic dipole braking, then $n \approx 3$ [156], while ≈ 5 if the predominant spin-down mechanism is GW braking [28, 157]. Assuming that the braking index is a constant and that the predominant GW mode is quadrupolar, one can integrate the decay equation to get an expression for the GW frequency evolution:

$$f_{\rm gw}(t) = f_0 \left(1 + \frac{t}{\tau}\right)^{\frac{1}{1-n}}. \tag{3.20}$$

Here f_0 is the initial GW frequency at some time $t = 0$ and τ is the spin-down timescale,

$$\tau = -\frac{(\pi f_0)^{1-n}}{k(1-n)}. \tag{3.21}$$

If the moment of inertia of the remnant is constant with time, the GW strain from

a rigid rotating body can be written as,

$$h_0 = \frac{4\pi^2 G}{c^4}\frac{I_{zz}\epsilon f_{\rm gw}^2}{d}, \tag{3.22}$$

where ϵ is the ellipticity, defined in Eq. 1.47, and d is the distance to the remnant from Earth. Inserting Eq. 3.20 we get,

$$h_0 = \frac{4\pi^2 G}{c^4}\frac{I_{zz}\epsilon f_0^2}{d}\left(1+\frac{t}{\tau}\right)^{\frac{2}{1-n}}. \tag{3.23}$$

The + and × polarization strains are then given by,

$$\begin{aligned} h_+ &= h_0\,\frac{1+\cos^2\iota}{2}\cos\left(2\pi f_{\rm gw}t\right) \\ h_\times &= h_0\,\cos\iota\sin\left(2\pi f_{\rm gw}t\right). \end{aligned} \tag{3.24}$$

Here, ι is the inclination angle, i.e the angle between the line of sight and the rotational axis of the remnant. One can further show that the total energy emitted through quadrupolar GW emission by time t is given by [155],

$$E_{\rm gw}(t) = -\frac{32\pi^6 G I_{zz}^2 f_0^6 \epsilon^2 \tau}{5c^5}\frac{n-1}{n-7}\left\{\left(1+\frac{t}{\tau}\right)^{\frac{7-n}{1-n}}-1\right\}. \tag{3.25}$$

Note that this model contains several assumptions that are known to break down eventually. The moment of inertia and the eccentricity are assumed to be constant; that is probably not a good assumption for rapidly spinning down remnants. Perhaps more importantly, the breaking index is assumed to be a constant, locking the mechanism to spin down. As a consequence, the expression for energy in Eq. 3.25 can become unphysical by exceeding the initial rotational kinetic energy of the remnant [155]. Moreover, the waveform completely neglects nuclear and thermal effects and phenomena like fall-back accretion. This means that the waveform described in Eq. 3.24 is very idealized and must be treated as more akin to a toy model. Chapter 4 deals with making robust inferences in the presence of such model uncertainties, focusing primarily on the phase of the waveform.

Despite these modeling deficiencies, the spin-down model captures the fundamentals

of the physics of spin down and will suffice to set limits for the algorithms considered here. Unmodeled search algorithms are robust to modeling uncertainties, but the flip side is they are also insensitive to model features.

3.6.2 Limits set following GW170817

In the STAMP search, a simulated signal campaign – hereafter referred to as injections – was undertaken using the spin-down model, and the Stochtrack algorithm for recovery to assess sensitivity and derive limits. The injections all used a distance of 40 Mpc and a braking index of $n = 5$ that assumes that the predominant mechanism for energy loss was through GW emission. The range of f_0 used was from 500-3000 Hz and of τ from $10^2 - 10^4$ Hz. The injection set used two inclinations angles of $\cos\iota = 1$ and $\cos\iota = 0$, corresponding to the best-case and worst-case inclination scenarios, respectively. The injections were done over the $\sim$ 8.5 days of H1 and L1 data after the coalescence of GW170817, but with an unphysical time shift.

Figure 3.6 shows the 90% limits on the energy radiated in the post-merger phase for τ = 100, 1000 and 10,000 s, assuming optimal inclination angle. The limits represent the energy that needs to be radiated by a spinning down remnant at a distance of 40 Mpc to generate a strain signal that, when marginalized over time, is statistically more significant than any on-source triggers 90 % of the time. As can be seen, the energy limits are much larger than the entire remnant's mass energy at all frequencies, making them quite unphysical. This is, unfortunately, a result of the relatively poor high frequency sensitivity of the current generation of detectors.

While the remnant's mass-energy represents an absolute physical limit on how much energy can be radiated, a more realistic energy limit for a spin-down model is its rotational kinetic energy. One can use this to convert the energy limits into limits on distance sensitivity that are shown in Fig. 3.7. These limits represent the closest distance before which a remnant, initially spinning at an an angular frequency of πf_0, would have been detected 90% of the time if it has radiated all of its rotational kinetic energy through GWs. A comparison of the limits with the remnant's actual distance of 40 Mpc immediately tells us how much we fell short at O2 sensitivity. With that definition, we see that we were about a factor of $\sim$ 40 away in sensitivity from being able to constrain GW radiation at frequencies close to 500 Hz.

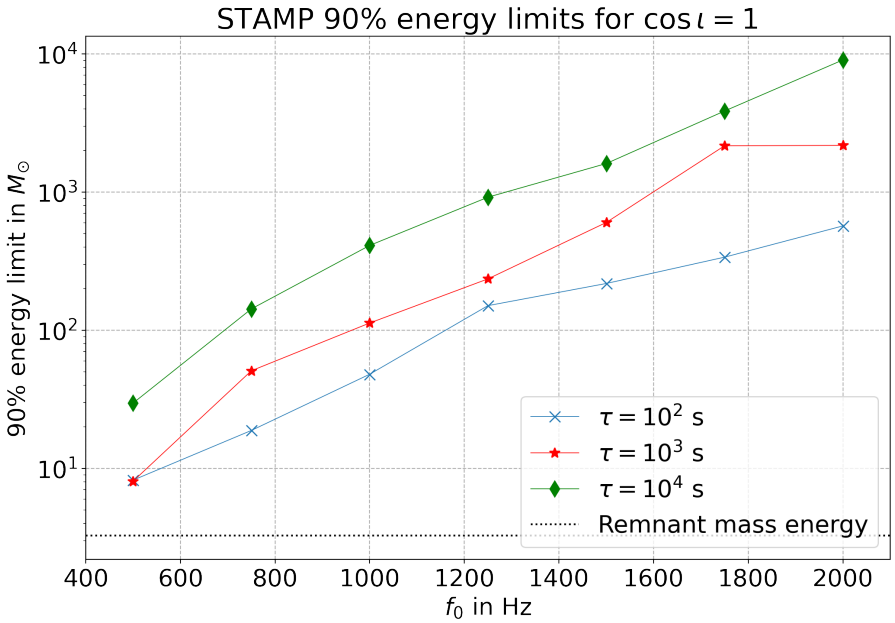


Figure 3.6: Optimal inclination limits on the energy radiated in the long-lived post-merger phase by the STAMP analysis are shown for different values of τ. The markers show the actual values derived from the injection analysis. All injections shown here used an $\iota = 1$ and $n = 5$. The dotted line represents the remnant's mass-energy, corresponding to $\approx 3.265 M_\odot$. The data used to generate this figure is from Ref. [3], reused here with permission.

This picture is consistent with the results of the other pipelines used in this analysis, as shown by Fig. 3.8. The figure shows the 90% limits of the various pipelines used in the search for $\tau = 100$ s. These pipelines make different choices in injection procedures, such as assumption about inclination angle and the data used for injections, and have different modes of analysis that makes an apples-to-apples comparison impossible. Nevertheless, the fact that they all roughly agree with each other indicates that the limits established are a product of the data itself, robust to differences in each pipeline's choices and assumptions.

A table with limits of all the injections done by the STAMP pipeline is included in Appendix. A.

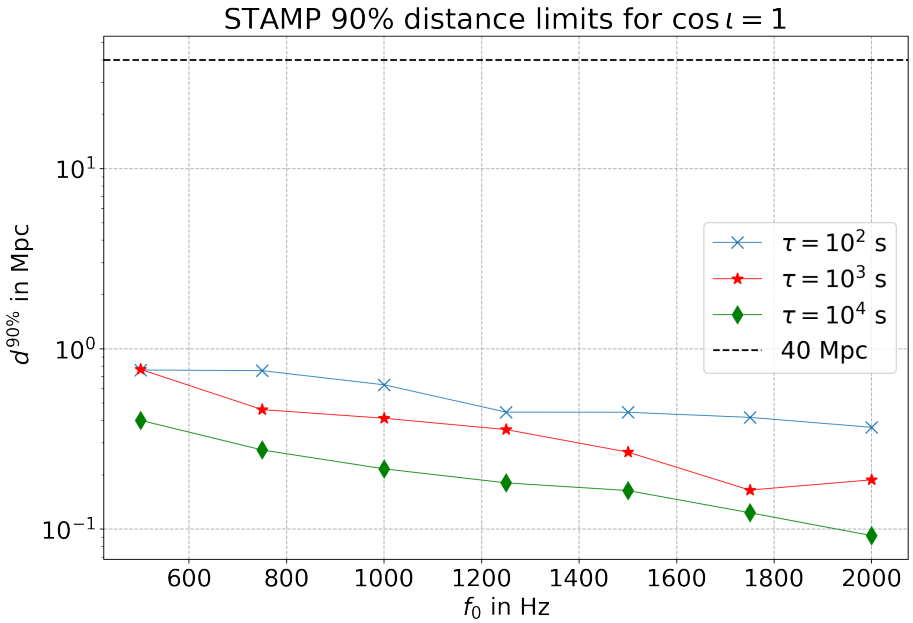


Figure 3.7: Optimal inclination limits on the distance to the merger remnant by the STAMP analysis are shown for different values of τ. The markers show the actual values derived from the injection analysis. All injections shown here used an $\iota = 1$ and $n = 5$. The dashed line shows the actual distance of the remnant at 40 Mpc. The data used to generate this figure is from Ref. [3], reused here with permission.

3.7 Discussion and conclusion

The results from the long-lived post-merger study [3] show that the network at O2 sensitivity could be sensitive to long-lived post-merger signals only at distances of less than 1 Mpc. If we did have such a close merger, we could constrain the fraction of energy lost through GW even in the absence of a detection.

An increase in detector sensitivities, especially at frequencies above 500 Hz, is necessary to be sensitive to remnants $\mathcal{O}(10)$ Mpc away like GW170817. This is consistent with the results of the other post-merger searches [10, 131] too, which saw that at least a factor of 10 increase in high-frequency sensitivity is necessary. Planned improvements of detector technologies like frequency-dependent quantum squeezing are expected to be extremely helpful in this regard.

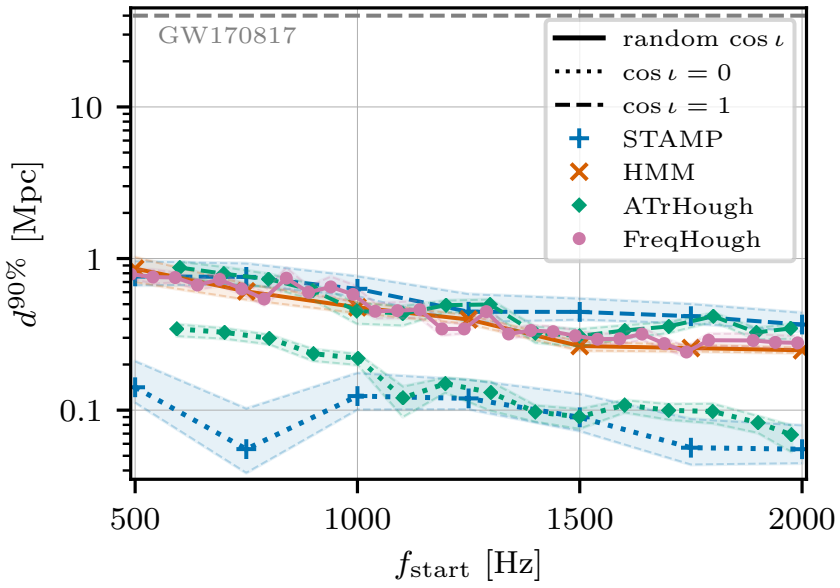


Figure 3.8: A comparison of the distance limits for the different pipelines in the long-lived post-merger remnant search for $\tau = 100$ s. Here FreqHough and AtrHough stand for the Generalized FrequencyHough and Adaptive Transient Hough pipelines respectively. The HMM and the FreqHough pipelines randomized over $\cos\iota$ while STAMP and AtrHough considered only the best and the worth case inclinations. The starting frequency f_0 is indicated by the f_{start} parameter. The shaded regions represent 90% confidence intervals on the limits. Figure reproduced from Ref. [3] with permission.

As such, post-merger signals are considered an important target for the next generation of ground-based detectors like Cosmic Explorer and Einstein Telescope [158, 159]. These detectors will have $\mathcal{O}(10)$ km long arms and constitute a significant improvement over current generation detectors throughout the frequency band. Detector designs have also been proposed, like the Australian Neutron Star Extreme Matter Observatory, which retain the current $\mathcal{O}(1)$ km arms but are specially tuned to the kiloHertz band. Exploring the post-merger phase is a key goal of such designs [160].

In conjunction with the detector improvements, one also needs better algorithms that can target the post-merger signals. One such attempt, focusing again on the long-lived post-merger signals, combined the HMM tracking algorithm (which was already used on auto-power TF-maps during the GW170817 post-merger search) with the cross-power SNR TF-maps of STAMP. This joins the strengths of both the methods and

yields about a factor of 2 increase in strain sensitivity to spin-down magnetar model waveforms, when compared to the seedless algorithm [151]. Another possibility is to combine information from multiple potential remnants to probe universal features of the underlying physics. Such improved algorithms will meet detector improvements halfway and can make detecting the post-merger phase a reality with the subsequent phases of ground-based detectors such as the third-generation detectors. It might then become possible to constrain the modes of emission of a newly-born NS – which has never been studied before – and even detect them through GWs. The STAMP analysis developed here, in conjunction with the other pipelines used, provide a foundation of such future searches and estimates the degree of improvement needed to get there.

Chapter 4

Methods for constraining the Gravitational-Wave Afterglow From a Binary Neutron Star Coalescence

This chapter develops Bayesian parameter estimation methods to constrain post-merger GW radiation from long-lived remnants. A novel phase agnostic likelihood model is developed that is robust to certain modeling uncertainties. The chapter follows a paper published in the Monthly Notices of the Royal Astronomical Society [11].

Abstract

Binary neutron star mergers are rich laboratories for physics, accessible with ground-based interferometric gravitational-wave detectors such as Advanced LIGO and Advanced Virgo. If a neutron star remnant survives the merger, it can emit gravitational waves that might be detectable with the current or next-generation detectors. Chapter 3 explored searches for such post-merger gravitational waves conducted following GW170817. The physics of the long-lived post-merger phase is not well understood and makes modeling difficult. In particular, the phase of the gravitational-wave signal is

not well modeled. This Chapter explores methods for using long-duration post-merger gravitational-wave signals to constrain the parameters and the properties of the remnant. A phase-agnostic likelihood model is developed that uses only the spectral content for parameter estimation and demonstrates the calculation of a Bayesian upper limit in the absence of a signal. With the millisecond magnetar model, it is shown that for an event like GW170817, the ellipticity of a long-lived remnant can be constrained to less than about 0.5 in the parameter space used.

4.1 Introduction

The detection of gravitational-wave signals from BBH mergers [23, 161, 81], and the BNS merger GW170817 [25] by Advanced LIGO and Advanced Virgo [24, 26] in the O1 and O2 runs show that compact binary coalescences are primary sources of GWs for terrestrial GW detectors. BNS mergers, in particular, provide an extremely rich environment for studying physics at conditions unattainable on Earth.

Searches by the LIGO scientific collaboration and the Virgo collaboration following GW170817 did not find any evidence for post-merger GWs from a neutron star remnant [131, 3, 10, 162] [1] as described in Ch. 3. The nature of the remnant of a BNS coalescence depends on the mass and spin of the remnant and the nuclear equation of state(for e.g., [122, 164]). One possible outcome is the formation of a rapidly rotating, highly magnetized, and long-lived ($t \geq 10$ s) massive neutron star (NS). Although no conclusive evidence for a long-lived remnant was found following GW170817, observations of X-ray afterglows of short gamma-ray bursts support this evolutionary pathway for a relatively large fraction of mergers [165]. Observations of GWs from a long-lived post-merger remnant could help probe the complex physics governing the pre and post-merger phase, as well as help constrain the equation of state of massive remnants. Some predictions of GW signals from long-lived remnants suggest they may be observable with second-generation observatories out to 40 Mpc (for e.g., [166]), although more realistic analyses that account for the energy budget [155] are more pessimistic and suggest that they might only be detectable with third-generation GW detectors [167, 168].

[1] There has been, however, a detection claim of a short-duration post-merger signal [163].

There has been much work in exploring the GW emission from newly born magnetars. The nature of these GWs can depend sensitively on a number of aspects of neutron-star physics, including early cooling before the transition to superfluidity, the effect of the magnetic field on the equilibrium shape, the internal dynamical state of a fully degenerate, oblique rotator, and the strength of the electromagnetic torque on the newly-born NS (e.g. [169, 166, 170, 171]). The amplitude and phase of the GWs depend on the complicated details of these physical mechanisms, which makes modeling and astrophysical inference from post-merger GW detections difficult. While there exist unmodeled Bayesian inference pipelines like Bayeswave [99, 172] - which can fit any signal using a wavelet expansion of variable dimensions - such analysis can be computationally expensive for long-transient signals considered in this chapter.

This chapter develops methods for Bayesian inference of long-transient signals, which are robust towards some modeling uncertainties. It focuses on the phase of the signal in particular and derives phase-agnostic likelihoods, which depend only on the spectral content of the signal. This likelihood is then used in the context of Bayesian parameter estimation to constrain intrinsic properties of a (long-lived) remnant such as ellipticity and the braking index, using the millisecond magnetar model waveform [173, 155] as an example waveform. Note that while this waveform model is used to study and demonstrate parameter-estimation methods, this is not to claim that it is a realistic model of long-lived post-merger emission. It is shown how this formalism performs both in the presence and the absence of a signal and how upper limits can be placed on gravitational-wave emission in the case of non-detection.

4.2 Millisecond Magnetar Model

The search for post-merger emission from GW170817 by LIGO and Virgo [131, 3] considered a variety of possible signals, ranging from sub-second to hour-long timescales. In particular, the search for a signal from a long-lived remnant was based on a model derived from the dynamics of spinning down neutron stars proposed by [155] and [173]. This model - here-after referred to as the millisecond magnetar model - derives the frequency evolution of the waveform from a spinning-down nascent neutron star with an arbitrary but fixed braking index n. Some of the details of the model are shown below.

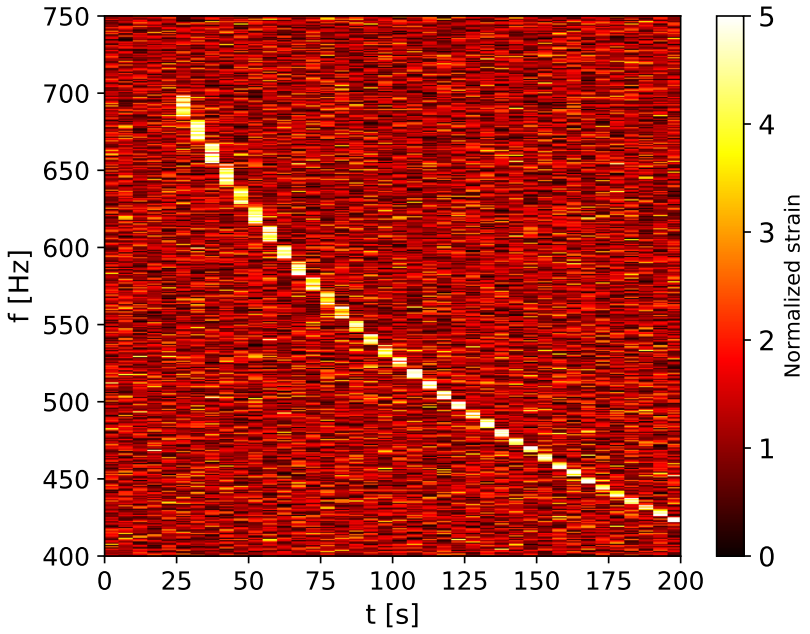


Figure 4.1: A normalized strain time-frequency map made with simulated Gaussian data recolored with O2 noise. A loud signal has been added for demonstration. The duration of each fast Fourier transform is 4 seconds and the entire map is 200 seconds long. Reproduced from Ref. [11] with permission.

We assume that the rotational evolution of the star is described by the torque equation: $\dot{\Omega} \propto \Omega^n$, where Ω is the star's angular frequency. We also assume quadrupole GW emission caused by a non-zero ellipticity of the neutron star, so that $f(t) = \Omega(t)/\pi$. Integrating the torque equation with a fixed braking index yields the GW frequency evolution:

$$f(t) = f_0 \left(1 + \frac{t - t_0}{\tau}\right)^{1/(1-n)}, \quad t \geq t_0. \tag{4.1}$$

Here t_0 is the start time of the emission (with some definition of $t = 0$), f_0 is the initial GW frequency (at $t = t_0$) and τ is the spin-down timescale. Equation 4.1 can describe emission from a variety of physical processes responsible for spin-down. For example, $n = 3$ describes magnetic dipole-powered spin-down in vacuum, while $n = 5$ describes spin-down powered by emission of quadrupolar gravitational waves. The amplitude of

the GW signal decreases with time as

$$h(t) = h_0 \left(1 + \frac{t - t_0}{\tau}\right)^{2/(1-n)}, \tag{4.2}$$

where we define an initial amplitude parameter h_0 as

$$h_0 = \frac{4\pi^2 G}{c^4} \frac{I_{zz}\epsilon}{d} f_0^2. \tag{4.3}$$

Here, d is the distance of the source, I_{zz} is the moment of inertia and ϵ is the ellipticity of the neutron star defined in Eq. 1.47.

4.3 Likelihood Model

A common way to search for GW sources that are difficult to model accurately is to look for excess power in time-frequency representations (TF-maps) of GW detector data [174, 175, 176, 132]. To detect GWs, the TF-maps are parsed by pattern-recognition algorithms looking for statistically significant clusters of pixels — for example, seeded [177, 142] and seedless [145, 146, 144] clustering algorithms using predefined templates have been widely used in the past. This was described in more detail in Sec. 3.3 and 3.4.

This chapter uses TF-maps of discrete (complex) Fourier transforms of the data, normalized by the noise power spectral density (PSD). Note that unlike the STAMP maps, these are single detector maps and no cross correlation is employed. An example map with a loud simulated signal is shown in Figure 4.1. We assume that the noise is Gaussian and stationary over the analysis period. We also assume that the noise in different pixels, i.e. across frequency and time, is uncorrelated. While this is a reasonable approximation for simulated advanced LIGO data used in this chapter (see Appendix. B), this might not always be true for real interferometric data, in which case alternate basis like Discrete wavelet transforms (for e.g., see [178]) could be a much more suitable choice.

We start with a likelihood model that assumes the residual noise when the signal is subtracted from the data is colored Gaussian noise. The Gaussian likelihood for a pixel

in TF-map is given by Eq. 2.8, reproduced below [95] [2] :

$$\mathcal{L}(\tilde{d}_{ij}|\boldsymbol{\theta}) = \frac{2}{\pi T S_j^n} \exp\left(-\frac{2}{T}\frac{|\tilde{d}_{ij} - \tilde{h}_{ij}(\boldsymbol{\theta})|^2}{S_j^n}\right), \tag{4.4}$$

where i, j are indices for the pixel at the i-th time-segment and j-th frequency bin of the TF-map. The terms $\tilde{d}_{ij}$, $\tilde{h}_{ij}$ and S_j^n are the Fourier transform of the data, the signal model, and the noise PSD of the pixel i, j respectively. The term T is the duration of the data used for the Fourier transform, and $\boldsymbol{\theta}$ is the vector of model parameters.

Given the uncertainty in the physics of the post-merger model describing the phase evolution of the remnants, we do not expect the signal model to be accurate. We need to incorporate our ignorance of the true phase of the signal when analyzing the data. One way to do this is by marginalizing the phase of each pixel independently of other pixels[3]. We begin by explicitly separating the phase term in Eq. 4.4,

$$\mathcal{L}(\tilde{d}_{ij}|\boldsymbol{\theta}, \phi_{ij}^h) = \frac{2}{\pi T S_j^n} \exp\left(-\frac{2}{T}\frac{|\tilde{d}_{ij}|^2 + |\tilde{h}_{ij}|^2}{S_j^n}\right) \exp\left(\frac{4}{T}\frac{|\tilde{d}_{ij}||\tilde{h}_{ij}|\cos(\phi_{ij}^d - \phi_{ij}^h)}{S_j^n}\right), \tag{4.5}$$

where ϕ_{ij}^h and ϕ_{ij}^d are the model and data phase in the pixel i, j. We marginalize over ϕ_{ij}^h (or equivalently over $\phi_{ij} = \phi_{ij}^d - \phi_{ij}^h$) with a uniform prior as a natural choice. The marginalization integral is,

$$I_{ij}^{\phi} = \frac{1}{2\pi}\int_0^{2\pi} d\phi \exp\left(\frac{4}{T}\frac{|\tilde{d}_{ij}||\tilde{h}_{ij}|}{S_j^n}\cos\phi\right). \tag{4.6}$$

The integral can be written in terms of a zeroth-order modified Bessel function of the first kind I_0 [179] so that the marginalized likelihood is,

$$\mathcal{L}_{\phi}(\tilde{d}_{ij}|\boldsymbol{\theta}) = \frac{2}{\pi T S_j^n} \exp\left(-\frac{2}{T}\frac{|\tilde{d}_{ij}|^2 + |\tilde{h}_{ij}|^2}{S_j^n}\right) I_0\left[\frac{4}{T}\frac{|\tilde{d}_{ij}||\tilde{h}_{ij}|}{S_j^n}\right]. \tag{4.7}$$

[2]Note that the correct normalization of the Gaussian likelihood function in the frequency domain should be proportional to σ^{-2}, and not to σ^{-1} like in real time-domain data. This is because frequency-domain noise is generally complex in which both the real and imaginary parts of the noise are independently Gaussian. See Appendix D of [93] for a careful examination of this.

[3]The phase marginalization being done here is different from the one used in parameter estimation analysis of compact binary coalescence, e.g., [95]. The phase evolution of compact binary waveforms is well understood, and it is only the initial phase that is marginalized over.

This is the likelihood for a single pixel of one interferometer. We take the product of likelihoods over all pixels to extend it over the entire TF-map. The simplest way to incorporate multiple detectors is to take the product of likelihoods for each detector (but see Sec 6.2):

$$\mathcal{L}(\{d_k\}|\boldsymbol{\theta}) = \prod_{i,j,k} \mathcal{L}_\phi(d_{ijk}|\boldsymbol{\theta}), \tag{4.8}$$

where k is an index over interferometers.

Tests of the likelihood in Eq. 4.7 show that the recovered parameters suffer from biases unless the exact PSD of the noise is known. A common way to estimate the noise PSD is by calculating the mean of the PSDs of neighboring or off-source data segments. This estimate has a variance about the actual PSD of the noise, which would need to be accounted for when large amounts of data are analyzed. One way to do this is to marginalize over the true PSD in a pixel given our measurement of S_j^n. Starting with the Gaussian likelihood in Eq. 4.4 and using a χ^2 prior for the true PSD gives a likelihood based on a Student-t distribution as described in Eq. 2.22. For pixel i, j we get:

$$\mathcal{L}_S(\tilde{d}_{ij}|\boldsymbol{\theta}) = \frac{4\,\Gamma\left(1+\frac{\nu}{2}\right)}{\pi T \nu S_j^n\,\Gamma(\nu/2)} \left[1 + \frac{4}{T}\frac{|\tilde{d}_{ij} - \tilde{h}_{ij}|^2}{\nu\, S_j^n}\right]^{-\left(1+\frac{\nu}{2}\right)}. \tag{4.9}$$

Here ν is the number of degrees of freedom of the χ^2 prior. A natural value for ν is $\nu = 2N$, where N is the number of data segments used to calculate S_j^n. As pointed out in [180], Student-t distributions with fewer degrees of freedom have larger tails, implying that they better account for uncertainties in the noise PSD and yield more robust inferences. This is, however, limited by the assumption that the noise is stationary. Methods that simultaneously model the noise along with the signal, such as in Refs. [181] and [99] would work better for non-stationary real data. Using fewer degrees of freedom is seen to give better inferences; we estimate the PSD using $N = 40$ segments and use $\nu = N$ in all the examples shown in this chapter.

Having accounted for the PSD variance, we marginalize over the phase of the pixels

in the signal TF-map. We define the α_{ij},β_{ij} and γ variables as

$$\begin{aligned}\alpha_{ij} =& \frac{2\,\Gamma\left(1+\frac{\nu}{2}\right)}{\pi\nu S_j^n\,\Gamma(\nu/2)}\left[\frac{\nu\,S_j^n + |\tilde{d}_{ij}|^2 + |\tilde{h}_{ij}|^2}{\nu\,S_j^n}\right]^{-\left(1+\frac{\nu}{2}\right)}, \\ \beta_{ij} =& \frac{2\,|d_{ij}||h_{ij}|}{|d_{ij}|^2 + |h_{ij}|^2 + \nu S_j^n}\,,\ \gamma = -\left(1+\frac{\nu}{2}\right).\end{aligned} \tag{4.10}$$

This allows us to write the Eq. 4.9 as,

$$\mathcal{L}_S^{ij}(\tilde{d}_{ij}|\bar{\theta}) = \alpha_{ij}\left[1-\beta_{ij}\cos(\phi_{ij}^s - \phi_{ij}^h)\right]^{\gamma}. \tag{4.11}$$

We now marginalize over the phase term $\phi_{ij} = \phi_{ij}^s - \phi_{ij}^h$ as before;

$$\mathcal{L}_S^{ij}(\tilde{d}_{ij}|\bar{\theta}) = \frac{\alpha_{ij}}{2\pi}\int_0^{2\pi} d\phi_{ij}\,[1-\beta_{ij}\cos\phi_{ij}]^{\gamma}. \tag{4.12}$$

This integral can be written in terms of Gauss hypergeometric functions ${}_2F_1$,

$$\begin{aligned}\mathcal{L}_{S,\phi}(\tilde{d}_{ij}|\bar{\theta}) = &\frac{\alpha_{ij}}{2}(1-\beta_{ij})^{\gamma}\,{}_2F_1\left(0.5,-\gamma,1,\frac{2\beta_{ij}}{\beta_{ij}-1}\right) + \\ &\frac{\alpha_{ij}}{2}(1+\beta_{ij})^{\gamma}\,{}_2F_1\left(0.5,-\gamma,1,\frac{2\beta_{ij}}{\beta_{ij}+1}\right),\end{aligned} \tag{4.13}$$

which gives the phase and PSD marginalized likelihood for each pixel.

4.4 Analysis

The likelihood in Eq. 4.13 is used to recover a simulated signal from the millisecond magnetar model added to Gaussian noise colored with the O2 PSD of Hanford and Livingston Advanced LIGO detectors. We make TF-maps that are 200 seconds long, divided into 4 s Tukey-windowed FFT pixels. In this analysis, we assume that we know the distance d and the sky-location of the remnant, which were simulated to be the same as GW170817 and its electromagnetic counterpart [25, 27], i.e (ra, decl) = (13.1634 Hrs, $-23.3185°$) and $d = 40$ Mpc. We also assume a polarization angle of ψ

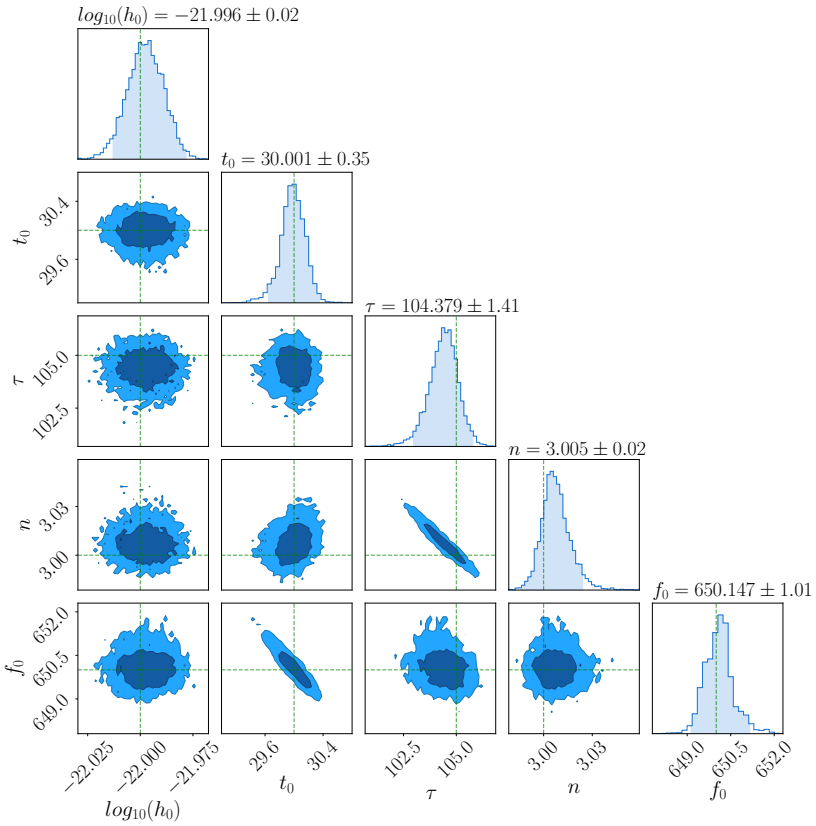


Figure 4.2: Posteriors for a millisecond magnetar model simulation based on Eq 4.1. The colored regions in the 1-d posteriors show 95% confidence intervals. The vertical green lines in the left panel are the true values corresponding to $log_{10}(h_0) = -22.0, t_0 = 30\,s, \tau = 105\,s, n = 3$ and $f_0 = 650$ Hz from left to right. The dark and the light regions in the 2-d posteriors are 68% and 95% confidence levels respectively. Reproduced from Ref. [11] with permission.

$= 0$ and an optimal orientation of the remnant, i.e. $\cos\iota = 1$. The polarization angle is the angle of rotation between the detector axes and the angle used for calculating the GW polarization in Eq. 3.24.

We sample over a five-dimensional parameter space $\boldsymbol{\theta} = \{h_0, t_0, \tau, n, f_0\}$ using PyMultiNest [112], a python wrapper for the Nested Sampling implementation of MultiNest [111]. Flat priors are used on all parameters [4] except h_0, for which a uniform

[4]The priors extend from $20s$ to $70s$ for t_0, from 2.3 to 5 on n, and from 625 Hz to 725 Hz on f_0. The parameter τ is degenerate with t_0, so in place of τ we actually sample over $T = \tau + t_0$ with a flat prior between 50s to 150s.

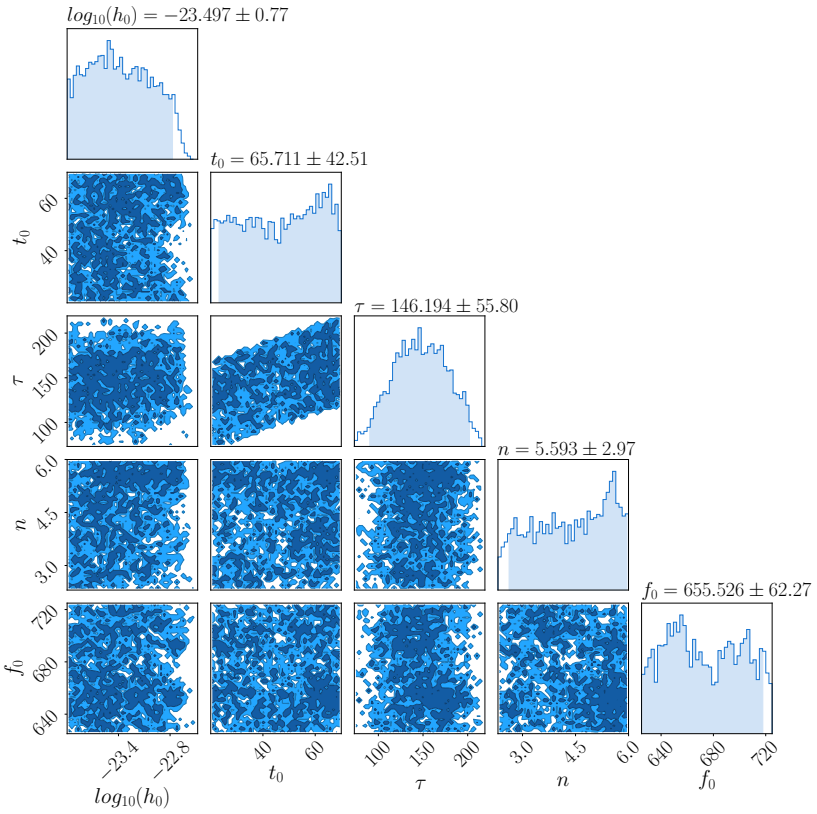


Figure 4.3: Posteriors for a millisecond magnetar model recovery over Gaussian noise with no signal. The colored regions in the 1-d posteriors show 95% confidence intervals. Reproduced from Ref. [11] with permission.

in log prior from 10^{-24} to 10^{-21} us used. Figure 4.2 shows an example of the parameter estimation of a loud simulated signal. In this case, the parameters are constrained roughly to a percent level.

Figure 4.3 shows results from an analysis with only Gaussian noise. In the absence of a signal, the posterior of the signal amplitude h_0 can be used to place upper limits on some properties of the remnant. Here, we get a 95% upper limit on h_0 of 2.1×10^{-23} with a uniform in log prior. Using the posterior samples and with Eq. 4.3, we can constrain the physical parameters of the remnants. In this case for example assuming a distance of 40 Mpc and the same fiducial moment of inertia as in Ref. [3] of $I_{zz} = 4.34 \times 10^{38}\,\mathrm{kg\,m^2}$, gives us a 95% limit on ellipticity of 0.499.

In reality, the distance and the sky-position might not be known in the absence of an

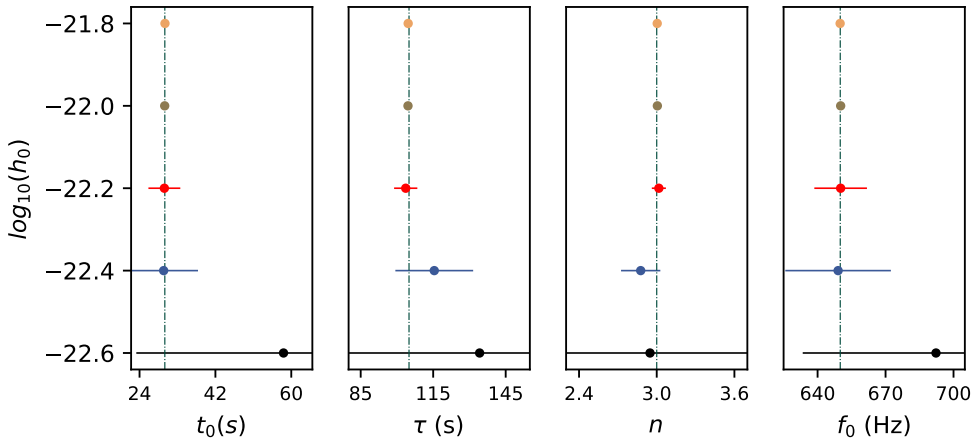


Figure 4.4: Posterior recoveries for the millisecond magnetar model showing the levels at which spectral parameters are constrained at different amplitude values. The vertical axis on the left shows the amplitude values used for the simulations. The solid dots are the maximum a posteriori values and the error bars correspond to 95% confidence levels. The vertical dashed-dotted line are the true values. Note that the x axis for these plots do not show to the full prior range and we have zoomed in to see the error bars better. Reproduced from Ref. [11] with permission.

electromagnetic counterpart, and the moment of inertia, polarization, and inclination angle of the remnant would also not be known precisely. These extra sources of uncertainty would need to be folded into both the analysis and the upper-limit calculation as needed, either as extra parameters or using constraints from other measurements (for e.g., one can use the distance measurement from the inspiral signal).

The upper limit on h_0 is also consistent with Figure 4.4, which attempts to recover the simulated signals of different amplitude levels while keeping constant the spectral parameters. The figure shows 95% confidence intervals with which the spectral parameters are recovered at different amplitudes. While the posteriors are well constrained for $h_0 \geq 4.0 \times 10^{-23}$, for a signal with amplitude $h_0 \leq 2.5 \times 10^{-23}$, the posteriors span almost the entire prior range.

The marginalized likelihood, Eq. 4.13 is finally used on simulated signals with phase evolution that is different from the model assumptions. A sanity test was first performed in the frequency domain with phase-scrambled maps — which are TF-maps with random fluctuations added to the phase of each pixel. We found that the recovered posteriors

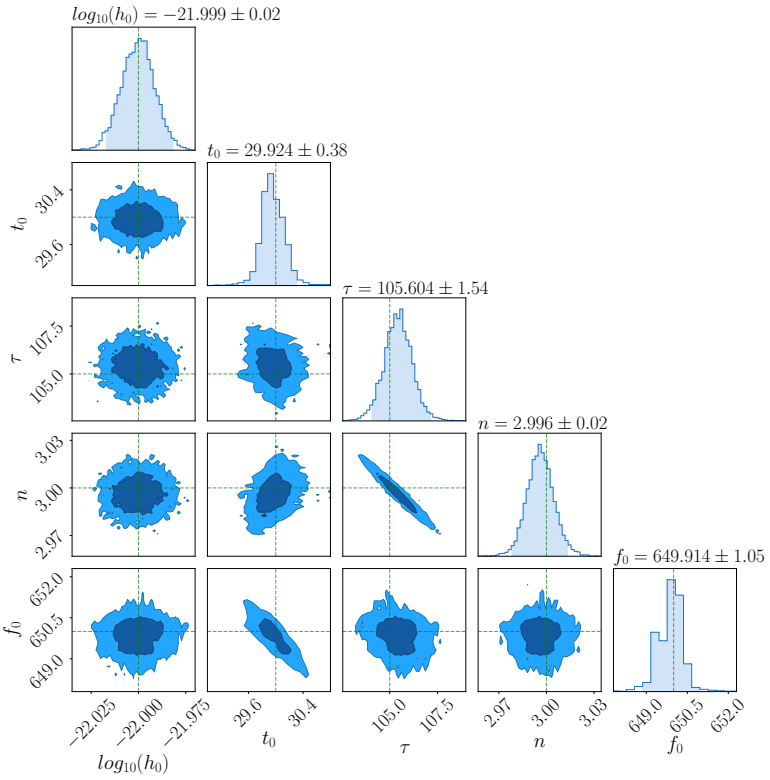


Figure 4.5: Posteriors recoveries for a simulation with small fluctuations added to the phase evolution. The colored regions in the 1-d posteriors show 95% confidence intervals. The vertical green lines in the left panel are the true values corresponding to $log_{10}(h_0) = -22.0, t_0 = 30\,s, \tau = 105\,s, n = 3$ and $f_0 = 650$ Hz from left to right. The dark and the light regions in the 2-d posteriors are 68% and 95% confidence levels respectively.

are consistent with the true values, as expected mathematically for the likelihood in Eq. 4.13 which only uses the absolute values of the signal model and data pixels.

A similar test was also performed by adding fluctuations to the time-domain phase evolution of the signal, thus mimicking a case where the phase evolution model was incorrect in a stochastic manner. Altering the waveform phase in the time domain also alters the spectral shape in the frequency domain. Care is needed to ensure that the time-domain fluctuations are not large enough to affect the frequency evolution of the signal too much. In the test performed here, a random phase error drawn from a Normal distribution with a standard deviation of $\approx 28.65°$ was added to the model phase at

each time step when generating the simulated signal. This test examines the robustness we get not just by phase marginalization but also the Student-t PSD marginalization. The Figure 4.5 show the results for a signal with small fluctuations added in the time domain, which demonstrate posterior recoveries consistent with the true parameters of the model.

4.5 Conclusion

Post-merger signals from neutron stars are a promising source of GWs for second and third-generation gravitational-wave detectors. This chapter describes the application of a Bayesian likelihood formalism to the characterization of long-duration post-merger signals from binary neutron star mergers. This formalism is demonstrated to be robust against fluctuations in phase evolution and is capable of constraining and measuring important astrophysical parameters like the spin, the braking index, the moment of inertia, and the ellipticity of magnetars.

The possibility of estimating the braking index of the remnant NS with gravitational-wave data is particularly noteworthy. There have been only two measurements of braking indices of millisecond magnetars to date using x-ray observations following short gamma-ray bursts [173]. Braking index measurements would be of particular interest since they give direct information on the underlying mechanics of the spin-down. In conjugation with ongoing developments [182, 183, 184] in modeling of post-merger GW emission, parameter estimation methods can also help constrain the nuclear equation of state at very high densities.

As the second-generation gravitational-wave detectors progress towards their design sensitivity, it is plausible that there will be a detection of a long-transient gravitational-wave signal in the coming observing runs. In addition to post-merger searches, analyses of these signals also benefit from the development of parameter estimations methods that make minimal model assumptions. Developments of sky localization methods robust to modeling are also planned as they can especially important for finding transients. One assumption made throughout this work is that the spectral model is well known. Parameter estimation methods that can handle a wider range of model uncertainties should also be developed to help in extracting as much astrophysical information from

future detections as possible in an unbiased way.

Chapter 5

Gravitational Waves as Probes of Cosmological Large Scale Structure

This chapter develops methods to use GW detections as tracers of the large-scale structure of the Universe and to cross correlate them with other tracers. In particular, it focuses on measuring angular two-point correlations as a simpler case, parameterized as an expansion in Legendre polynomials. Simulated data is then used to test the methods developed. The chapter follows a paper published in the Physical Review D journal [12].

Abstract

The aLIGO and aVirgo detectors have detected ten binary black hole mergers by the end of their second observing run. These mergers have already allowed constraints to be placed on the population distribution of black holes in the Universe, which will only improve with more detections and increasing sensitivity of the detectors. This chapter develops statistical techniques to measure the angular distribution of black hole mergers by measuring their statistical N-point correlations through hierarchical Bayesian inference and applies it to the special case of two-point angular correlations using a Legendre polynomial basis on the sky. Building on the mixture model formalism

introduced in Ref. [185], methods are developed to measure two-point correlations with no threshold on significance, allowing us to target the ensemble of sub-threshold BBH mergers not resolvable with the current generation of ground-based detectors. These methods can also be used to correlate GWs with other probes of large-scale angular structure like galaxy counts and validate both techniques through simulations.

5.1 Introduction

The direct detections of GWs by the aLIGO and aVirgo detectors [23, 25, 81] have given us a new tool to probe the Universe. GWs can carry astrophysical and cosmological information not accessible through electromagnetic observations. This is especially true for BBH mergers which otherwise, leave no electromagnetic trace [1]. The aLIGO and aVirgo detectors have detected ten BBH mergers [81] in the first two observing runs, with many more candidate events in the recently completed third observing run (O3) [187]. These detections have allowed us to constrain the mass, spin, and redshift distributions of BBH progenitors, along with measuring the rate of mergers in the local Universe [115]. Additional events have also been claimed by groups analyzing publicly available data from O1 and O2 [188, 189, 190]. In addition to these population properties, recently, there have also been studies measuring the angular distributions of the BBH merger events in the published GWTC-1 catalog from O1 and O2 [191, 192].

With the current generation of ground-based GW detectors, there are many more such mergers of stellar origin which are individually unresolvable. Finding the combined signal from this ensemble of unresolved mergers has traditionally been a key target of stochastic GW searches [193, 194, 93] that employ cross correlation between detectors to detect an astrophysical background. Multiple analyses have been developed to measure anisotropies in the stochastic background as applied to ground-based GW detector data [195, 196, 197, 198, 199]. In recent years, there have also been theoretical predictions about the anisotropic properties of astrophysical BBH backgrounds [200, 201, 202, 203, 204, 205, 206, 207, 208] which suggest that the lowest order multipoles — traditionally thought to be the ones stochastic searches are most sensitive

[1] However there have been a few claims of potential electromagnetic counterparts to binary black hole mergers, see [186, 84]

to — are at least an order of magnitude smaller than the monopole, placing them beyond the reach of current generation detectors using standard cross correlation based methods. Moreover, because of the relatively small number of BBH merger events that occur in an observation time, shot noise from statistical Poisson effects can dominate over the astrophysical contribution to the higher-order multipoles, making their discovery and measurement further difficult [209, 210].

This chapter develops methods to probe the statistical properties of the angular distribution of the ensemble of BBH mergers, folding in the discrete nature of events. Two different results are being described here. The first is a general way to measure the angular N-point correlations of the field of mergers through hierarchical Bayesian inference, focusing on the special case of two-point correlations. This method is then applied to simulations in the context of the mixture model framework developed by Smith and Thrane [185].

The mixture model approach places no thresholds on the signal-to-noise ratio (SNR) of the events and can jointly draw inferences from well-resolved events and events in the astrophysical background (the so-called sub-threshold events) accounting for their discrete nature, thus removing the somewhat artificial distinction between classical resolvable compact binary coalescence (CBC) signals and stochastic GW backgrounds from CBC sources. The mixture model framework promises in general, to be much more sensitive than the cross-correlation search in probing the astrophysical background since it looks for very specific signals based on CBC waveform models, and with the addition of hierarchical Bayesian inference can also be used to estimate properties of the population of BBH systems [211] in addition to the rate of their mergers. This extra sensitivity is promising in attempting to detect the angular structure of the ensemble of binary mergers.

The rest of the chapter is organized as follows. Section 5.2 defines the N-point correlation function and using that as a correlated prior, we write down an expression for the Bayesian signal evidence. Section 5.3 uses the general expression for the case of two-point correlations between BBH mergers and also shows how a similar expression for evidence can be written for cross-correlating BBH mergers with a different tracer of the large scale structure like galaxy counts. Building upon the mixture model from Ref. [185], Sec. 5.4 shows how the methods from the previous sections can be used for

Bayesian inference of anisotropies in an ensemble of BBH mergers. Section 5.5 describes the simulations and presents results for the two-point correlation method for both BBH-BBH and BBH-galaxy count correlations, followed by a discussion in Sec. 5.6 of potential data quality issues that could arise when these techniques are applied to real data.

5.2 BBH N-point Correlation

Statistical N-point correlation functions have been traditionally very useful in measuring the statistical clustering of galaxies [212], with the most important statistic perhaps being the two-point correlation function. Since BBH mergers are discrete events, we can correlate multiple such observed events to measure the statistical properties of the field of such mergers, in particular, to probe the angular structure of the ensemble of mergers and their progenitors. The tool of choice for such *population inference* is hierarchical Bayesian modeling, in which a suitably modeled prior contains the population parameters to be measured. The reader is directed to Ref. [213] for an overview of Bayesian inference in the field of GW analysis, including hierarchical inference, and to Ref. [214] for a guide to aLIGO-aVirgo noise and data.

Given a GW data segment d, the Bayesian evidence or the marginal likelihood that it contains a GW signal is given by:

$$\mathcal{Z}^0 = \int d\vec{\lambda}\, \mathcal{L}(d|\vec{\lambda})\, \pi^0(\vec{\lambda}), \tag{5.1}$$

where $\vec{\lambda}$ includes all relevant parameters needed for describing BBH merger waveforms like masses, spins, distance, sky position etc. The term $\mathcal{L}(d|\vec{\lambda})$ is the standard likelihood used in GW searches based on the noise statistics of colored, frequency-domain Gaussian noise [95, 96] described in Eq. 2.12 , and $\pi^0(\vec{\lambda})$ are some standard fiducial priors on $\vec{\lambda}$. Note that the parameters describing the sky position $\hat{\Omega} = (\theta, \phi)$ are part of $\vec{\lambda}$.

Measuring the N-th statistical correlation requires N data segments each containing a signal. Firstly, under the standard assumption that they are *i.i.d* events drawn from the same population and using the fiducial priors, the evidence for the hypothesis that all of them contain a BBH signal follows from Eq. 5.1 as,

$$\mathcal{Z}_N^0 = \int \prod_i^N \left[d\vec{\lambda}_i \, \mathcal{L}(d_i|\vec{\lambda}_i) \, \pi^0(\vec{\lambda}_i) \right], \tag{5.2}$$

where i is an index over data segments. It has been assumed here that the probability for a data segment to contain more than a single BBH merger is negligible. Similarly, the probability of a signal cutting across segments is assumed to be negligible. The duration T of the data segment will have to be chosen to ensure that this is a good assumption. The typical durations used are between $4 - 16$ seconds.

To measure statistical correlations in the population distribution, we use a joint correlated prior between multiple segments. To write the evidence for the new N-point angular correlations hypothesis, we replace the fiducial angular prior in Eq. 5.2 with a prior correlated across the N data segments, while retaining the fiducial priors on all other parameters, viz [2]:

$$\mathcal{Z}_N^c(\vec{\Lambda}) = \int \left[\zeta^N(\{\hat{\Omega}_j\}|\vec{\Lambda}) \prod_{i=1}^N \left(d\vec{\lambda}_i \, \frac{\mathcal{L}(d_i|\vec{\lambda}_i)\pi^0(\vec{\lambda}_i)}{\pi^0(\hat{\Omega}_i)} \right) \right], \tag{5.3}$$

where $\pi^0(\hat{\Omega}_i)$ are the fiducial *uncorrelated* priors on angular parameters of each segment. The term $\{\hat{\Omega}_j\}$ refers to the set of all the directional parameters for the N segments, with the N-point correlation function $\zeta^N(\{\hat{\Omega}_i\}|\vec{\Lambda})$ being the joint prior on them with hyperparameters $\vec{\Lambda}$. The hyperparameters describe the population distribution of the mergers, in this case the angular distribution. The correlation function is interpreted as the probability density distribution of N objects being at the angular positions of $\{\hat{\Omega}_j\}$ depending on the choice of $\vec{\Lambda}$ [212].

It is important to note that under the correlated hypothesis, the individual BBH mergers are still treated as independent events – as they must be since they are spatially and temporally separated – which allows the likelihoods in Eq. 5.3 to be multiplied with each other. The statistical correlations only describe the population distribution of the BBH mergers, and they are correlations in probabilities.

[2] In this chapter $\mathcal{Z}_N^c(\vec{\Lambda})$ is described as the Bayesian evidence. It could alternatively be called the marginal likelihood on $\vec{\Lambda}$ marginalized over all the intrinsic parameters of the mergers, as if often used for such hierarchical inference problems.

Equation 5.3 above casts the problem as one of joint inference of both hyperparameters $\vec{\Lambda}$ and parameters of the BBH merger $\vec{\lambda}_i$. The posterior distributions are usually calculated by means of nested or Markov chain Monte Carlo samplers. Generally, however, it is not necessary to redo the sampling at every point in the hyperparameter space. Instead the posterior samples generated while using the fiducial choice of priors on the BBH parameters $\vec{\lambda}_i$ can be *recycled* to provide inferences in the hyperparameter space [213]. To see this in the case of the directional parameters, we first use Bayes theorem on Eq. 5.3 to get:

$$\mathcal{Z}_N^c(\vec{\Lambda}) = \int \left[\mathcal{Z}_N^0 \, \zeta^N(\{\hat{\Omega}_j\}|\vec{\Lambda}) \prod_{i=1}^{N} \left(d\vec{\lambda}_i \, \frac{P^0(\vec{\lambda}_i|d_i)}{\pi^0(\hat{\Omega}_i)} \right) \right], \tag{5.4}$$

where $P^0(\vec{\lambda}_i|d_i)$ is the fiducial posterior for segment i obtained using prior $\pi^0(\vec{\lambda}_i)$. The evidence $\mathcal{Z}_N^0$ (defined in Eq. 5.2) comes up as the normalization factor when using Bayes theorem and can be pulled out of the integral. We then marginalize over all the non-directional parameters to get:

$$\mathcal{Z}_N^c(\vec{\Lambda}) = \mathcal{Z}_N^0 \int \left[\zeta^N(\{\hat{\Omega}_j\}|\vec{\Lambda}) \prod_{i}^{N} \left(d^2\hat{\Omega}_i \, \frac{P^0(\hat{\Omega}_i|d_i)}{\pi^0(\hat{\Omega}_i)} \right) \right]. \tag{5.5}$$

The integrand is then just the expectation value of the ratio of the new correlated prior and the fiducial angular priors. When we have posterior samples rather than a continuous measurement of $P^0(\hat{\Omega}_i|d_i)$, as as described in Sec. 2.5.1, this can be approximated as,

$$\mathcal{Z}_N^c(\vec{\Lambda}) \approx \mathcal{Z}_N^0 \sum \frac{\zeta(\{\hat{\Omega}_i^{m^i}\} \,|\, \vec{\Lambda})}{\prod_i M_i \, \pi^0(\hat{\Omega}_i^{m^i})}. \tag{5.6}$$

Here i is an index over segments while $m^i = 1...M_i$ is an index over the posterior samples of that segment, and M_i is the number of posterior samples for that segment. The summation is over all possible N-point correlations between posterior samples across segments the total number of which is given by $\prod_i M_i$. For the directional parameters $\hat{\Omega}_i$, the fiducial angular priors are usually taken to be isotropic for each segment i.e $\pi^0(\hat{\Omega}_i) = 1/4\pi$.

5.3 Two-Point correlation

5.3.1 BBH Two-Point Correlation

When probing large-scale cosmological structure, one is usually more interested in the statistical properties of the distribution rather than the specific realization in our Universe. Theoretical models also only predict the statistical properties. Statistical isotropy is a standard assumption made when measuring cosmological correlations, which often simplifies Bayesian searches by reducing the number of parameters needed for modeling structures of a given angular scale. With these assumptions, measuring the two-point correlation of the BBH background is the most interesting and the simplest case of N-point correlations. This is mainly because of the fact that fluctuations in dark matter density and galaxy distribution can be traced back to Gaussian fluctuations in the early Universe in linear scales. The two-point correlations are a measure of the variance of the Gaussian fluctuations. Two-point correlation-based priors can also be used in a straightforward manner for directly correlating GW data with EM probes of structure. Hence in this section and for the remainder of the chapter, we focus on the specific case of $N = 2$ to apply the N-point correlation formalism.

Assuming statistical isotropy and homogeneity, we define the two-point correlation function as the probability density of two BBH mergers i and j being at an angular separation Δ_{ij} [212, 215]. Under these assumptions, the correlation function ζ can be expanded in the basis of Legendre polynomials $\mathcal{P}_\ell(\Delta_{ij})$ with coefficients $\{C_\ell\}$ as parameters:

$$\zeta(\Delta_{ij}^{mn} \,|\, \{C_\ell\}) = \frac{1}{(4\pi)^2} \sum_\ell (2\ell + 1)\, C_\ell\, \mathcal{P}_\ell(\cos(\Delta_{ij}^{mn})), \tag{5.7}$$

where the factor of $(4\pi)^2$ is required to normalize ζ as a probability distribution function. The convention adopted here is to define ζ as the complete two-point correlation function instead of the over and under densities as is usually done in galaxy number count analyses. Our definition can be converted into the latter by simply writing the isotropic (monopole) term separately. If $\hat{\Omega}_i = (\theta_i, \phi_i)$ and $\hat{\Omega}_j = (\theta_j, \phi_j)$ are the coordinates of the BBH merger events, the angular separation between them on the two sphere Δ_{ij} is given by:

$$\cos(\Delta_{ij}) = \sin\theta_i \sin\theta_j \cos(\phi_i - \phi_j) + \cos\theta_i \cos\theta_j. \tag{5.8}$$

The joint evidence $\mathcal{Z}^c_{ij}$ for the two-point correlation hypothesis as a function of $\{C_\ell\}$ then follows from Eq. 5.3:

$$\mathcal{Z}^c_{ij}(\{C_\ell\}) = \int d\vec{\lambda}_i\, d\vec{\lambda}_j\, \pi^0(\vec{\lambda}_i)\, \pi^0(\vec{\lambda}_j) \frac{\mathcal{L}(d_i|\vec{\lambda}_i)\mathcal{L}(d_j|\vec{\lambda}_j)}{\pi^0(\hat{\Omega}_i)\pi^0(\hat{\Omega}_j)}\, \zeta(\Delta_{ij}|\{C_\ell\}), \tag{5.9}$$

with the equivalent of Eq. 5.6 given by,

$$\mathcal{Z}^c_{ij}(\{C_\ell\}) \approx \frac{\mathcal{Z}^0_i \mathcal{Z}^0_j}{M_i M_j} \sum_m \sum_n \frac{\zeta(\Delta^{mn}_{ij} \,|\, \{C_\ell\})}{\pi^0(\hat{\Omega}^m_i)\pi^0(\hat{\Omega}^n_j)}. \tag{5.10}$$

Here $\mathcal{Z}^0_i$ and $\mathcal{Z}^0_j$ are the signal evidences using the fiducial angular priors for event i and j. The term Δ^{mn}_{ij} is the angular separation between the m-th sample and the n-th sample in the posteriors of the i and j data segments respectively. Finally using Eq. 5.7 and the isotropic value for $\pi_o(\hat{\Omega}_i) = 1/4\pi$ we get,

$$\mathcal{Z}^c_{ij}(\{C_\ell\}) \approx \frac{\mathcal{Z}^0_i \mathcal{Z}^0_j}{M_i M_j} \sum_{m,n} \sum_\ell (2\ell+1)\, C_\ell\, \mathcal{P}_\ell\left(\cos(\Delta^{mn}_{ij})\right). \tag{5.11}$$

5.3.2 BBH - Galaxy two-point correlation

If the progenitors of BBH mergers are black holes of stellar origin, we expect that their angular distribution will follow that of the large scale structure on the sky. Cross correlating this distribution with other tracers of structure like galaxy counts could allow us to probe this common matter distribution, and also test theories of structure and evolution. In this section, we show how two-point correlation evidence can also be written for measuring cross correlations between GW and galaxy distribution as seen by surveys like SDSS [216]. The theoretical prediction of the cross correlation will depend on details of redshift evolution of the star formation rate, among other things. However, we can again write it in a fairly model-independent way as an expansion in Legendre polynomials, assuming statistical isotropy.

In the case of cross correlation we now define the two-point correlation function $\zeta_{ij}(\Delta_{ij}|\{C_\ell\})$ as the probability of having a BBH merger i at an angular separation of

Δ_{ij} from a galaxy j [3]. Since the positions of the galaxies are usually known to very high precision compared to that of GW sources, we assume that the uncertainty associated with them is negligible. The two-point correlation prior can be written in the Legendre polynomial basis as:

$$\zeta(\Delta_{ij}|\{C_\ell\}) = \frac{1}{4\pi}\sum_\ell (2\ell+1)\, C_\ell\, \mathcal{P}_\ell(\Delta_{ij}) \tag{5.12}$$

Note that the prefactor here is different from Eq. 5.7 because here we just have just one angular integral when normalizing; over that of the BBH merger $\hat{\Omega}_i$ as compared to two angular integrals in the latter. The evidence when data segment i is "correlated" with galaxy j is then given by,

$$\mathcal{Z}^g_{ij}(\{C_\ell\}) = \int d\vec{\lambda}_i\, \pi^0(\vec{\lambda}_i)\, \frac{\mathcal{L}(d_i|\vec{\lambda}_i)}{\pi^0(\hat{\Omega}_i)}\, \zeta(\Delta_{ij}|\{C_\ell\}), \tag{5.13}$$

where $\vec{\lambda}_i$ are as before all the BBH parameters and $\hat{\Omega}_i$ are the directional parameters. Incorporating sample recycling this can be approximated to,

$$\mathcal{Z}^g_{ij}(\{C_\ell\}) \approx \frac{\mathcal{Z}^0_i}{M_i}\sum_n \frac{\zeta(\Delta^n_{ij}|\{C_\ell\})}{\pi^0(\hat{\Omega}^n_i)} \tag{5.14}$$

The term Δ^n_{ij} is the angular separation between the n-th sample in the posteriors of the i-th data-segment and the j-th galaxy. Finally using Eq. 5.12 and $\pi_o(\hat{\Omega}_i) = 1/4\pi$, the evidence becomes,

$$\mathcal{Z}^g_{ij}(\{C_\ell\}) \approx \frac{\mathcal{Z}^0_i}{M_i}\sum_n \sum_\ell (2\ell+1) C_\ell \mathcal{P}_\ell(\Delta_{ij}) \tag{5.15}$$

5.4 Mixture model formalism

The Eqs. 5.11 and 5.15 for two-point correlations derived in the previous sections are generally valid for measuring anisotropies with any kinds of transient GW data. Along

[3] There are two distinct but related questions we can ask. The first is the probability of having a BBH event at a given angular separation, while the second is having a BBH progenitor at a separation with respect to a galaxy. We focus on the former here because it is simpler, but the two would be related by a Poisson distribution of the rate of BBH mergers in the Universe.

with Eq. 5.6 which is also similarly general, they constitute the first important result of this chapter. Such correlation techniques can be used with the various catalogs of events after accounting for selection effects as has been done with other hierarchical analyses (see for example [191, 217, 218]). However, for the rest of this chapter, the correlation functions are applied in the context of the mixture model analysis developed in Ref. [185, 211].

The mixture model formalism works by using compact binary coalescence parameter estimation algorithms on many available data segments without any cutoff on significance or SNR, allowing us to dig deep into the background of sub-threshold events. While this removes biases due to selection effects, we need to account for the fact that only some, a priori unknown, fraction of the segments (referred to as the signal duty cycle) will contain a real astrophysical signal. The analysis then uses Bayesian signal and noise evidence from these data segments to construct posterior probability distributions for the signal duty cycle, as well as for the desired population hyperparameters. Some important details are reproduced here.

We divide the data into segments of duration T, chosen such that it is much larger than the inspiral time scale of BBH mergers in the aLIGO-aVirgo frequency band, while also being much smaller than the inverse rate of BBH mergers in the Universe. A choice of $\tau = 4\text{s}$ - 16s sits comfortably within this range. Under the assumption that there are no non-Gaussian glitches in the data, two possible hypotheses exist for each data segment:

1. There is a BBH signal in the data segment.

2. There is only instrumental Gaussian noise in the data segment.

We denote by ξ_S the signal duty cycle, i.e the fraction of data segments which contain a BBH merger signal. With just two hypotheses, the noise duty cycle is then $\xi_N = 1 - \xi_S$. We then construct a mixture model likelihood for ξ_S for the data segment i using the signal evidence $\mathcal{Z}_S^i(\vec{\Lambda})$ and noise evidence $\mathcal{Z}_N^i$,

$$\mathcal{L}(d_i|\xi_S, \vec{\Lambda}) = \xi_S \, \mathcal{Z}_S^i(\vec{\Lambda}) + \xi_N \mathcal{Z}_N^i. \tag{5.16}$$

The noise evidence $\mathcal{Z}_N^i$ is just the likelihood that the data d_i comprises only of instrumental colored Gaussian noise. The signal evidence depends on population parameters $\vec{\Lambda}$ modeled by priors $\pi(\vec{\lambda}|\vec{\Lambda})$ where $\vec{\lambda}$ are the intrinsic parameters for each event:

$$\mathcal{Z}_S^i(\vec{\Lambda}) = \int d\vec{\lambda}\, \mathcal{L}(d_i|\vec{\lambda})\, \pi(\vec{\lambda}|\vec{\Lambda}). \tag{5.17}$$

we get the posterior distribution for the duty cycle ξ_S and hyperparameters $\vec{\Lambda}$ using priors $\pi(\xi_S)$ and $\pi(\vec{\Lambda})$ respectively :

$$P(\xi_S, \vec{\Lambda}|d_i) = \left(\xi_S \mathcal{Z}_S^i(\vec{\Lambda}) + \xi_N \mathcal{Z}_N^i\right) \pi(\xi_S)\, \pi(\vec{\Lambda}). \tag{5.18}$$

Applying this formalism to the BBH-galaxy two-point correlations is straightforward. Using $\mathcal{Z}_{ij}^g(\{C_\ell\})$ defined in Eq. 5.15, and assuming that all the other population parameters have either been marginalized over, or are perfectly known, the posterior for correlating data segment i with galaxy j can be written as

$$P(\xi_S, C_\ell|d_i, g_j) = \left(\xi_S \mathcal{Z}_{ij}^g(\{C_\ell\}) + \xi_N \mathcal{Z}_N^i\right) \pi(\xi_S)\pi(\{C_\ell\}). \tag{5.19}$$

There are more hypothesis to consider when we apply the mixture model to BBH two-point correlations. For any two data segments i and j there are four hypotheses at play.

1. Both data segments have a signal: The evidence for this hypothesis is $\mathcal{Z}_{ij}^c$ calculated in Eq. 5.11 or Eq. 5.10 more generally.

2. Data segment i has a signal while data segment j has only noise: The evidence for this hypothesis is $\mathcal{Z}_i^0 \mathcal{Z}_j^N$ where $\mathcal{Z}_i^0$ is the signal evidence calculated using the fiducial isotropic prior.

3. Data segment j has a signal while data segment i has only noise: The evidence for this hypothesis $\mathcal{Z}_j^0 \mathcal{Z}_i^N$.

4. Both data segments have only noise: The evidence for this hypothesis is $\mathcal{Z}_i^N \mathcal{Z}_j^N$.

The joint mixture model likelihood for correlation between GW events i and j is then given by,

$$\mathcal{L}(d_i, d_j|\xi_S, \{C_\ell\}) = \xi_S^2 \mathcal{Z}_{ij}^c(\{C_\ell\}) + \xi_N^2 \mathcal{Z}_j^N \mathcal{Z}_i^N + \xi_S \xi_N (\mathcal{Z}_i^0 \mathcal{Z}_j^N + \mathcal{Z}_j^0 \mathcal{Z}_i^N) \tag{5.20}$$

Some care is needed when extending this to multiple data segments. Naively one might expect that two-point correlations between any two possible BBH pairs will have some extra information to be extracted. But one also needs to ensure that contradictory hypotheses are not mixed up. For example, suppose that we combine likelihoods for correlations over pairs $i-j$ and $j-k$. Then the hypothesis that both $i-j$ have a BBH merger signal is clearly incompatible with the hypothesis that both $j-k$ have only noise since they share a common data segment. The simplest way out of this is to multiply likelihoods only over independent pairs of data segments [4]. The posterior for multiple segments is then:

$$P(\xi_S, C_\ell|d_i, d_j) = \prod_{i,j} \left[\xi_S^2 \, \mathcal{Z}_{ij}^c(\{C_\ell\}) \, + \xi_N^2 \mathcal{Z}_j^N \, \mathcal{Z}_i^N \, + \xi_S \xi_N \left(\mathcal{Z}_i^0 \mathcal{Z}_j^N + \mathcal{Z}_j^0 \mathcal{Z}_i^N\right)\right] \pi(\xi_S)\pi(\{C_\ell\}) \tag{5.21}$$

A similar argument applies for galaxy-BBH correlations in Eq. 5.19. If we correlate a BBH merger with multiple galaxies we run the risk of multiplying contradictory hypotheses, which means we have to correlate data segments and galaxies in a one-on-one manner. Thus when extending this to multiple galaxies and data segments we again need to take products over independent pairs:

$$P(\xi_S, C_\ell|d_i, g_j) = \prod_{i,j} \left(\xi_S \mathcal{Z}_{ij}^g(\{C_\ell\}) + \xi_N \mathcal{Z}_N^i\right) \pi(\xi_S)\pi(\{C_\ell\}). \tag{5.22}$$

[4] The number of possible pairs can be very big; with N segments, the number of pairs grows as $\mathcal{O}(N^2)$. We argue that any randomly chosen possible pairing is statistically valid. A heuristic argument for this is each pair-wise correlation can be thought of as a random sample from the underlying probability distribution of the correlation function. We then wish to choose a subset of the correlations to represent the distribution which is valid if the method of choosing is random, and is independent of the actual values of the correlations. It follows that any such randomly chosen set of pairs should represent the same underlying pdf to within statistical fluctuations. This is the simplest method we found, but it is possible that it does not make the optimal usage of all the information available. If a better scheme exists, its discovery is left to the future.

5.5 Simulations and Recovery

We simulated the GW data by first generating a large number of BBH signals using the IMRPhenomPv2 waveforms [219, 220] distributed isotropically over the sky between 0.5 Gpc to 5 Gpc in luminosity distance uniform in comoving volume and in 4s segments. The signals were then added without any overlap to simulated aLIGO and aVirgo design sensitivity instrumental noise. We then ran a CBC parameter estimation algorithm over each segment using the same waveform to get posteriors and evidences using fiducial isotropic angular priors. We also ran the parameter estimation algorithm over segments that contained only simulated instrumental noise. Both the simulations and the parameter estimation were done using the BILBY pipeline [96], with the nested sampling package DYNESTY [110] used for the latter.

From this large database of segments and posteriors, we generate anisotropic simulations with desired values of $\{C_\ell\}$ and ξ_S by probabilistically choosing segments based on the true sky position of the signal. To do this, we pixelize the sky with Healpix [221, 222] and calculate a probability map on the sky by drawing from a multivariate Gaussian distribution. We calculate the mean and the covariance matrix of the Gaussian using the chosen values of $\{C_\ell\}$. The monopole i.e. C_0 gives the mean of the multivariate Gaussian. The covariance matrix can be computed by calculating the two-point correlation between pixels using the higher multipoles, dipole, and above. The probability map, along with the desired signal duty cycle ξ_S dictate the number of BBH events in each pixel, which are then randomly chosen from the previously generated database of simulated signals.

Fig. 5.1 show an example probability map generated with this method. Since the Legendre expansion describes a real field, it is possible that some of the pixels will have negative probability values. Such pixels are excised by setting their probabilities to zero so that no BBH mergers or galaxies are assigned to them. As we go towards smaller multipole moments relative to the monopole, this problem is expected to disappear. Any simulations made through $\{C_\ell\}$ will be susceptible to two kinds of noise. One is variance due to a specific realization of the map; this is similar to cosmic variance. The second is Poisson shot noise in the pixel. In order to correct for the noise effects and the excision of pixels we compute the $\{C_\ell\}$ values of the maps once they are made, and

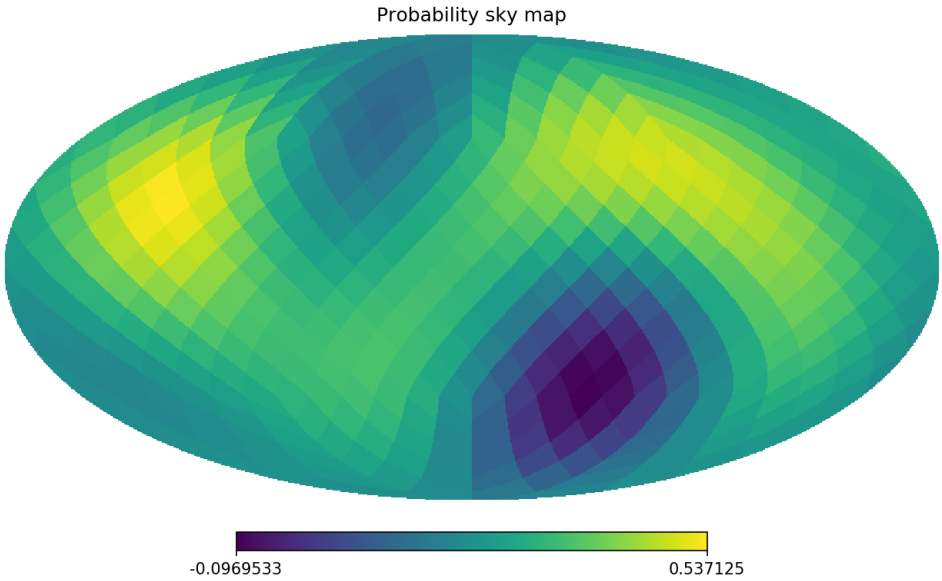


Figure 5.1: An example Mollweide map of the probability distribution on the sky generated by using the method described in Sec. 5.5, with an $\ell_{max} = 3$ and with $C_1 = 0.13, C_2 = 0.11$ and $C_3 = 0.11$. Some of the pixels have an unphysical negative probability; no black holes or galaxies are allocated to those pixels in the simulations. Reproduced from Ref. [12] with permission.

use those as the true values. While this is a simplistic solution, a more sophisticated correlation function modeling the noise effects could also be used to account for them.

All simulations shown in this chapter consist of 4s segments for aLIGO Hanford, aLIGO Livingston, and aVirgo interferometers. The random fraction of segments that contain a signal is given by the duty cycle value, chosen to be $\xi_S = 0.7$ for all simulations, with the rest being just Gaussian instrumental noise. The duty cycle value is chosen for computational reasons and is very large compared to realistic astrophysical rates. Instead, as a metric, we will use the effective time scales of the simulations, defined here as the amount of real data needed to have the same number of BBHs as in the simulation, assuming an average rate of 1 BBH every 4 minutes. For the 2.5×10^4, 4s long segments used for BBH two-point correlations this implies an effective time scale is ~ 48 days with $\xi_S = 0.7$. For the BBH-galaxy correlations with 2.2×10^4 segments, this gives a time scale of ~ 42 days.

Recovery corner plots from analyzing the BBH simulations with the two-point correlation method described in Sec. 5.3.1 are shown in Fig. 5.2. We assume that the expression could be cut of at $\ell_{max} = 3$, which is the same value used in generating the simulation. All higher multipole moments are set to zero. The $\{C_\ell\}$ describe the statistical correlations at different angular scales. The corner plots demonstrate that the statistical properties of the background are well recovered by the methods described here.

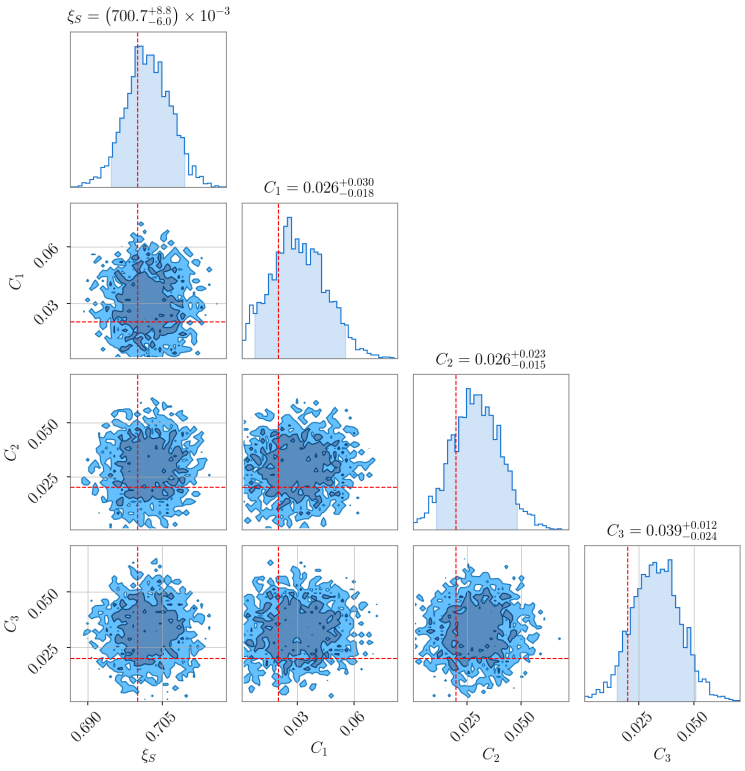


Figure 5.2: Plot showing the posterior distributions for the angular correlations $\{C_\ell\}$, and the duty cycle factor ξ_S for BBH-BBH two-point correlations with 2.5×10^4 data segments and $\ell_{max} = 3$. The monopole term is not an explicit parameter since it is normalized over and the other $\{C_\ell\}$ are normalized against it. The dashed red lines are the true values of the injected parameters with $\xi_S = 0.7$ which corresponds to 17.5×10^3 BBH signals, and $(C_1, C_2, C_3) = (0.018, 0.016, 0.019)$. We use uniform priors on both ξ_S and $\{C_\ell\}$; 0 to 1 on the former and 0 to 0.1 on the latter. The shaded regions in the 1-d posteriors correspond to symmetric 90% confidence intervals. Figure is reproduced from Ref. [12] with permission.

For the case of BBH-Galaxy correlation, we generated simultaneous simulations of BBH signals and galaxy counts. The BBH simulations were done in the same way as before, while a simulated map of galaxy positions was made through rejection sampling using the same probability map made for the GW case. We then measure the two-point correlation function by correlating posteriors of the GW data set with a mock all-sky galaxy catalog using the methods described in Sec. 5.3.2. Recovery plots from this analysis are shown in Fig. 5.3. Since galaxies have negligible uncertainty in sky position, the correlation allows us to probe deeper into the common statistical distribution of galaxies and BBH progenitors. In addition to smaller anisotropy values, the posteriors also demonstrate recovery of higher-order anisotropies by successfully recovering Legendre coefficients with $\ell_{max} = 5$ with a smaller amount of GW data.

5.6 Application to real data

5.6.1 Glitch Hypothesis

While this chapter relies only on simulations made in stationary Gaussian data, a brief discussion of data quality is in order to access applicability to real data. As pointed out in Ref. [185], handling non-Gaussian artifacts in GW detectors (called glitches) requires us to introduce additional hypotheses for each segment. A conservative assumption is used that glitches look like single detector BBH signals. The Bayesian evidence that there is a glitch in a segment in detector 1 is then just the evidence for a single detector signal hypothesis i.e,

$$Z^{g}_{(1)} \equiv Z^{S}_{(1)} = \int d\vec{\lambda}_{(1)} P(\vec{\lambda})_{(1)}) \, \pi(\vec{\lambda}_{(1)}). \tag{5.23}$$

The subscript here is an index over detectors while $\vec{\lambda}$ consist of all the BBH parameters as before. In Ref. [185] individual glitch hypotheses are constructed for each detector and are used to measure the glitch duty cycles for each individual detector. We simplify that somewhat by constructing a single catch-all hypothesis that there is a glitch at any one of the detectors in a data segment. For this hypothesis we rely on the assumption that it is unlikely for a glitch to occur along with a signal in a segment, and that it is also unlikely for glitches to occur in two or more detectors in the same segment. Under

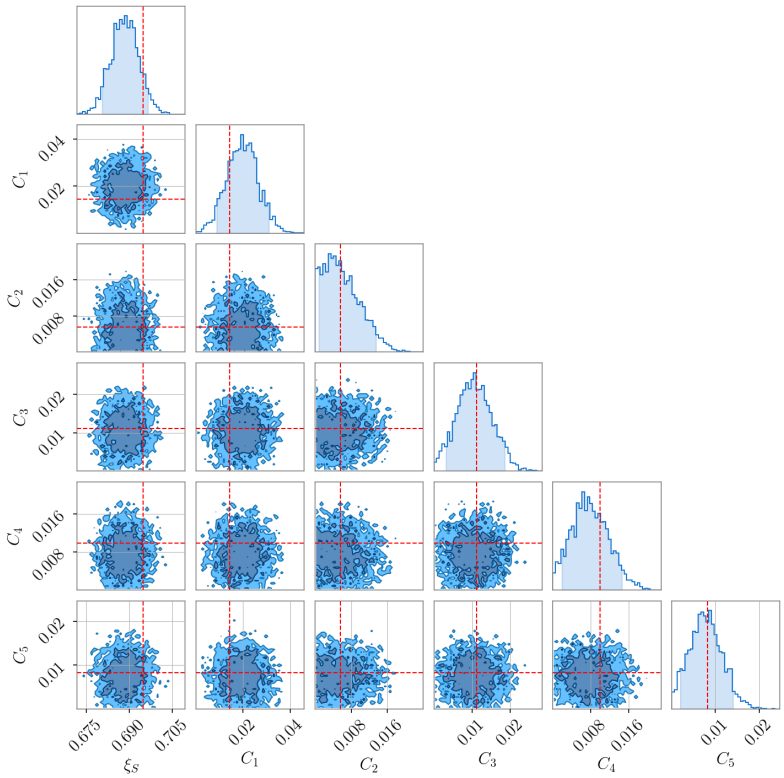


Figure 5.3: Plot showing the recovered $\{C_\ell\}$ as well as the duty cycle factor ξ_S using BBH-Galaxy two-point correlations with 2.2×10^4 data segments and $\ell_{max} = 5$. The monopole term is not an explicit parameter since it is normalized over and all other C_ℓ are normalized against it. We use uniform priors on both ξ_S and $\{C_\ell\}$; 0 to 1 on the former and 0 to 0.1 on the latter. The shaded regions in the 1-d posteriors correspond to symmetric 90% confidence intervals. The dashed red lines are the true values of the injected parameters with $\xi_S = 0.7$ which corresponds to 15.4×10^3 BBH signals and $(C_1, C_2, C_3, C_4, C_5) = (0.014, 0.006, 0.011, 0.01, 0.008)$. Figure is reproduced from Ref. [12] with permission.

these assumptions the glitch evidence for a segment for the case of three detectors is:

$$\mathcal{Z}^g = Z^N_{(1)} Z^N_{(2)} Z^S_{(3)} + Z^N_{(3)} Z^N_{(1)} Z^S_{(2)} + Z^N_{(2)} Z^N_{(3)} Z^S_{(1)} \tag{5.24}$$

The equivalent of Eq. 5.21 then becomes

$$\begin{aligned}\mathcal{L}(d_i, d_j|\xi_S, \xi_g, \{C_\ell\}) = \prod_{i,j} \Big[& \xi_S^2 \, \mathcal{Z}_{ij}^c(\{C_\ell\}) + + \xi_N^2 \mathcal{Z}_j^N \mathcal{Z}_i^N + \\ & \xi_S \xi_N \left(\mathcal{Z}_i^0 \mathcal{Z}_j^N + \mathcal{Z}_j^0 \mathcal{Z}_i^N \right) + \\ & \xi_g \xi_N \left(\mathcal{Z}_i^g \mathcal{Z}_j^N + \mathcal{Z}_j^g \mathcal{Z}_i^N \right) + \\ & \xi_g \xi_S \left(\mathcal{Z}_i^g \mathcal{Z}_j^0 + \mathcal{Z}_j^g \mathcal{Z}_i^0 \right) \Big],\end{aligned} \tag{5.25}$$

where ξ_g is the glitch duty cycle i.e the fraction of segments containing a glitch in one of the detectors. The duty cycle factors are now related as $\xi_S + \xi_N + \xi_g = 1$. We note again that this depends on coincident glitches between detectors being unlikely and additional data quality cuts might be required to ensure this requirement is met with real data. We will defer application of the N-point correlation methods to real data to a future work.

5.6.2 Application to real galaxy catalogs

The simplistic galaxy simulation in this chapter assumes that we can measure the galaxy field across the entire sky with equal sensitivity, which is not true for real galaxy surveys, especially because of obstruction from the dust and gas of the Milky Way. This effect is usually modeled by assuming that the observed field is filtered through a window function which captures the incompleteness of the observed galaxy distribution across the sky and has the effect of changing the spherical harmonic (and hence the multipole) expansion of the galaxy distribution (see for eg [223]). This would need to be accounted for when correlating with a real galaxy catalog.

We also point out that the two-point cross-correlation evidence described in Eq. 5.14 assumes a one-to-one pairing between GW data segments and galaxies. Since reusing them is not possible, we are forced to have the same number of galaxies as we have segments. Real galaxy catalogs will, of course, have tens or hundreds of millions of galaxies at the very least. One way to apply this formalism to cross correlating with real catalogs would be to randomly sample from them. For example, if one is working with a million GW data segments, one can randomly pick a million galaxies from a

catalog like SDSS and correlate them one on one with the GW data segments.

5.6.3 Sensitivity

While the broad localization of events in the GW posterior distributions is a major source of uncertainty for detecting and measuring anisotropies, another important source of noise arises from the Poisson statistics of the events. Under this shot noise, we would expect the uncertainty in measurements of spherical harmonic coefficients $a_{\ell m}$'s to fall as $1/\sqrt{N}$, where N are the number of events. Likewise, uncertainty in measurements of C_ℓ's should scale as $1/N$. If the two-point correlation method is applied to a catalog of GW events, then to measure a dipole anisotropy of $C_1 \sim 0.01$ we would need $\mathcal{O}(100)$ events to overcome the shot noise floor. This is broadly consistent with a simulated analysis done by Ref. [192], albeit in the context of next-generation detectors.

When the two-point correlation analysis is applied to sub-threshold events, as is done in this chapter, predicting sensitivity becomes more complicated. While the shot noise remains unchanged, an accurate estimate of the sensitivity would need simulations based on astrophysical realistic duty-cycles and population distributions, along with glitch rates of the detectors. For a distance cut off of $5\,Gpc$ and with a realistic duty cycle of 4×10^{-4}, the required time of detection of an isotropic signal was estimated to be ≈ 20 hours in [185]. Assuming the same shot noise-based scaling as above, we can then estimate that it would take $\mathcal{O}(100)$ days of data to detect $C_1 \sim 0.01$.

5.7 Conclusion

In this chapter, we have developed ways to measure the statistical N-point correlations of the angular distribution of BBH mergers, with emphasis on the specific case of two-point correlations. We have also shown how the two-point correlation method can be used to cross correlate BBH distribution with other tracers of large scale structure. Using the formalism developed in [185] and [211] we have demonstrated measurement of anisotropies on simulated data using two-point correlations. This method holds promise to delve deeper into the noise floor than standard stochastic searches and to measure anisotropies in the ensemble of binary mergers. The formalism can be extended to measure higher-order multipoles, too, if so desired.

Recently there have been studies on correlating GW data with the distribution of galaxies [224, 225, 226], and in particular on correlating the anisotropic stochastic maps from aLIGO-aVirgo with galaxy counts [227]. The methods developed in this chapter could provide a boost to such efforts. Theoretical modeling of the stochastic background from stellar mergers also suggests that BBH-galaxy correlations would be less susceptible to Poisson noise when measuring anisotropies than the GW side alone due to the relatively small number of BBH events [228, 229]. Finally, galaxy-BBH correlations could, in principle, allow us to probe differences in the relative distribution of galaxies and progenitors of GW. But this would perhaps require angular resolutions much smaller than possible with the current generation of detectors.

There are several ways to extend or apply the formalism developed in this work. One can apply the angular two-point correlation method to the catalog of events already published accounting for selection effects. The Bayesian posteriors also let us access the distance measurements of the events, so one can also consider measuring correlations in three dimensions rather than just over the two-sphere. This would give us the ability to directly measure the three-dimensional structure of matter and constrain the power spectrum of BBH progenitors through GW. The binary neutron star merger GW170817 demonstrated an application of GW towards cosmology through a GW measurement of the Hubble constant [8]. A similar idea was also recently explored in Ref. [230] in the context of third-generation detectors assuming a Gaussian localization of sources. Measurement of the power spectrum of matter distribution will enlarge the scope of GW as a tool for cosmological inference. The extension to three-dimensional two-point correlations can also be applied to catalogs of events to probe local structure. These will be explored in the future with application to GW catalogs and assessing the sensitivity of both the current generation of detectors and the next generation, which will have deeper redshift reach and more precise localization.

Chapter 6

Mapping the millihertz stochastic gravitational-wave confusion noise with LISA

This chapter develops Bayesian methods to map the GW sky using LISA, adopting a spherical harmonic basis. A novel decomposition based on Clebsch-Gordan coefficients is developed to make this an optimal mapping that is suitable for Bayesian inference. The chapter follows a paper that is available on arxiv [13] and has been submitted to the Monthly Notices of the Royal Astronomical Society. The BLIP pipeline used in this chapter has also been predominantly developed by this author.

Abstract

The millihertz gravitational-wave frequency band is expected to contain a rich symphony of signals with sources ranging from galactic white dwarf binaries to extreme mass ratio inspirals. Many of these gravitational-wave signals will not be individually resolvable. Instead, they will incoherently add to produce stochastic gravitational-wave confusion noise whose frequency content will be governed by the dynamics of the sources. The angular structure of the power of the confusion noise will be modulated by the distribution of the sources across the sky. Measurement of this structure can yield important

information about the distribution of sources on galactic and extra-galactic scales, their astrophysics, and their evolution over cosmic timescales. Moreover, since the confusion noise is part of the noise budget of LISA, mapping it will also be essential for studying resolvable signals. This chapter presents a Bayesian algorithm to probe the angular distribution of the stochastic gravitational-wave confusion noise with LISA using a spherical harmonic basis. We develop a technique based on Clebsch–Gordan coefficients to mathematically constrain the spherical harmonics to yield a non-negative distribution, making them optimal for expanding the gravitational-wave power and amenable to Bayesian inference. These techniques are then demonstrated using a series of simulations and analyses, including recovery of simulated distributed and localized sources of gravitational-wave power. We also apply this method to map the gravitational-wave foreground from galactic white-dwarfs using a simplified model of the galactic white dwarf distribution.

6.1 Introduction

The upcoming space-based Laser Interferometer Space Antenna (LISA) [231] promises access to the millihertz GW frequency band, which is inaccessible to terrestrial ground-based detectors like LIGO, Virgo and KAGRA. A rich collection of galactic and extra-galactic sources emit GWs at these frequencies. Among them are double white dwarfs (DWDs) from both within the Milky Way [232, 233] and neighboring satellite galaxies [234, 235], extreme mass ratio inspirals [236, 237, 238], supermassive blackhole binaries [239, 240], extragalactic stellar-mass BBH and BNS inspirals [241], and even exoplanets orbiting white dwarfs [242, 243]. Cosmological sources like cosmic strings [244], cosmological phase transitions [245] and primordial GW backgrounds [246] from the early Universe are also potentially accessible by LISA, not to mention the prospect of multi-wavelength GW science in conjunction with next-generation ground-based detectors (see for example Ref. [247])

The GWs from many astrophysical sources will not be individually detectable and will overlap to form a confusion noise in the detector, usually called the stochastic gravitational-wave background [248, 93]. The stochastic GW background from binary

inspirals is a major scientific target for both current ground-based detectors like Advanced LIGO [24] and Advanced Virgo [26], and for pulsar timing arrays [249, 250, 251]. The current limits from ground-based detectors are $\Omega_{\rm gw}(f = 25\ {\rm Hz}) < 4.8 \times 10^{-8}$ [252] for an isotropic background and $h_0 < (3.6\text{–}4.7) \times 10^{-25}$ for point sources [197]. The latest results from pulsar timing arrays are from the NANOGrav collaboration which see strong evidence for a common-spectrum red-noise process with a median strain of 1.92×10^{-15} at a frequency of 1 yr^{-1} [253].

In the case of LISA, there is also a GW foreground from galactic DWDs that stands above the instrumental noise for a part of the LISA band while still being stochastic in nature. While this foreground is considered an inconvenient noise source for resolvable signals, it is of astrophysical interest in its own right and contains useful information about the physical and spectral distribution of the DWDs [254, 255]. For the rest of the chapter, we will use the term stochastic gravitational-wave confusion noise (SGCN) to refer to both stochastic backgrounds and foregrounds collectively.

The angular structure of the GW power from a SGCN will directly follow the distribution of the sources which generates it. The antenna patterns of the detectors are not isotropic and also change throughout LISA's orbit, which ensures that the angular structure can, in principle, be measured with enough integration time. Spherical harmonic functions are a natural basis to describe the distribution of power on the sky and have been frequently used in algorithms developed to measure these anisotropies, both for LIGO and LISA, albeit usually in a frequentist manner [195, 256, 257, 258, 259, 260]. In particular, a frequentist maximum likelihood method to map GW power with LISA was recently developed in Ref. [261].

This chapter presents a Bayesian algorithm to map the power of an SGCN using a spherical harmonic basis. There are several advantages to developing a Bayesian version of this method, especially in the case of LISA, where the galactic foreground dominates. First, the Bayesian version can be better integrated with global analyses designed to extract multiple resolvable signals [262], in order to map the foreground simultaneously along with them. Accounting for the foreground in this way can be crucial for accurately inferring the properties of the resolvable signals. Frequentist searches also generally require the inversion of a Fisher matrix connecting different sky directions or harmonics. The poor angular sensitivity of GW detectors, LISA included, creates degeneracies that

make this inversion mathematically ill-conditioned, necessitating the use of techniques like singular value decomposition (see for example Refs. [195, 261]). These degeneracies are better accommodated with a Bayesian approach, which requires no such inversion. Additionally, the angular sensitivity of the spherical harmonic expansion is set by cutting off the expansion in ℓ at some $\ell_{\rm max}$ parameter value. This parameter value is chosen in a somewhat ad hoc way in frequentist searches but can be much more naturally accommodated in Bayesian searches by allowing the data to determine it.

Historically, one hindrance of a Bayesian spherical harmonic implementation has been that the generic expansion describes a complex field on the sky, while GW power is real and non-negative by definition. We demonstrate a way to mathematically impose this constraint using Clebsch–Gordan coefficients in a Bayesian spherical harmonic analysis. This mathematical technique was used recently to measure the angular distribution of GW detections by LIGO-Virgo [192], while a similar method was also recently used in the pulsar timing array band [263]. We also introduce the Bayesian LISA Pipeline (BLIP) designed to simulate LISA data, perform the spherical harmonic analysis on the simulated data, and conduct Bayesian inference to recover the simulated parameters.

The rest of the chapter is structured as follows. Section 6.2 reviews SGCNs, the spherical harmonic basis and the detector response function of LISA to GW power in the spherical harmonic basis. In Sec. 6.3 the Clebsch–Gordan decomposition for non-negative fields is calculated, and the parameterization necessary for Bayesian inference is developed. Section 6.4 introduces the BLIP pipeline and discusses the likelihood function and the configuration used for the analyses in this chapter. Section 6.5 demonstrates measurement of anisotropies in simulated LISA data in the spherical harmonic basis with the Clebsch–Gordan decomposition. Section 6.6 discusses applications of the technique to the galactic foreground followed by a discussion and conclusion in Sec. 6.7.

6.2 Stochastic gravitational-wave confusion noise

Astrophysical SGCNs result from an incoherent superposition of GWs from many disparate sources which are not individually resolvable [248] [1]. Appealing to the central

[1] which implies that the strength of the background depends on the sensitivity of the detector.

limit theorem, one can characterize the SGCN as colored Gaussian noise in the detectors [194]. This is usually a good assumption for sources in the LISA band that overlap with one other. The metric perturbation at $(t, \mathbf{x})$ corresponding to an SGCN can be written in the Fourier basis as,

$$h_{ij}(t,\mathbf{x}) = \sum_A \int_{-\infty}^{\infty} df \int d^2n\, \tilde{h}_A(f,\hat{\mathbf{n}})\, e^A_{ij}(\hat{\mathbf{n}}) \times \exp\{-2\pi i f(t - \hat{\mathbf{n}}\cdot\mathbf{x}/c)\}, \tag{6.1}$$

where $A = \{+, \times\}$ denotes polarization, c is the speed of light, $\hat{\mathbf{n}}$ is the directional unit vector and $e^A_{ij}(\hat{\mathbf{n}})$ are the polarization tensors [2]. The frequency of the SGCN is represented by f and $\tilde{h}_A(f,\hat{\mathbf{n}})$ are the Fourier components of the perturbations which satisfy the following,

$$\begin{aligned} \langle \tilde{h}_A(f,\hat{\mathbf{n}})\rangle &= 0, \\ \langle \tilde{h}_A(f,\hat{\mathbf{n}})\, \tilde{h}^*_{A'}(f',\hat{\mathbf{n}}')\rangle &= \frac{1}{2}\, S_A(f,\hat{\mathbf{n}})\, \delta_{A,A'}\, \delta(f-f')\, \delta^2(\hat{\mathbf{n}},\hat{\mathbf{n}}'). \end{aligned} \tag{6.2}$$

Under the assumption of Gaussianity the power spectrum $S_A(f,\hat{\mathbf{n}})$ is the main measurable quantity of an SGCN. For the rest of this chapter we will also assume that the SGCN is unpolarized i.e, $S_+(f,\hat{\mathbf{n}}) = S_\times(f,\hat{\mathbf{n}}) = 1/2\, S_{\rm gw}(f,\hat{\mathbf{n}})$. The power spectrum is conventionally characterized by the dimensionless energy density $\Omega_{\rm gw}(f,\hat{\mathbf{n}})$ per logarithmic frequency bin [194], related to $S_{\rm gw}(f,\hat{\mathbf{n}})$ via,

$$\Omega_{\rm gw}(f,\hat{\mathbf{n}}) = \frac{2\pi^2 f^3}{3H_0^2} S_{\rm gw}(f,\hat{\mathbf{n}}), \tag{6.3}$$

where H_0 is the Hubble constant. In general, the distribution of the power on the sky will not be isotropic but will rather trace the distribution of its sources. To describe the angular structure of the SGCN, it is assumed that its frequency and directional

[2] A point on notation is appropriate here. Generally, most paper use subscripts of ab rather than ij for the perturbation and polarization tensors. Conventionally $a, b \in \{1, 2\}$ to signify the transverse nature of the GWs, in a coordinate system where $\hat{\mathbf{n}}$ is aligned with the z axis. This is not a suitable choice for an SGCN which is an incoherent superposition of GWs from all directions. So we use the subscript $i, j \in \{1, 2, 3\}$ to clarify this point. See Sec. 7.8 of [28] for a more detailed description.

dependence can be factorized as

$$\Omega_{\rm gw}(f,\hat{\mathbf{n}}) = \Omega(f)\mathcal{P}(\hat{\mathbf{n}}). \tag{6.4}$$

The spectral shape of the SGCN is given by $\Omega(f)$ while $\mathcal{P}(\hat{\mathbf{n}})$ describes the angular distribution of the background, normalized so that:

$$\int_{\mathcal{S}^2} d^2n\, \mathcal{P}(\hat{\mathbf{n}}) = 1. \tag{6.5}$$

A common way to parameterize the spectral shape $\Omega(f)$ is to write it as a power law:

$$\Omega(f) = \Omega_0 \left(\frac{f}{f_0}\right)^{\alpha}, \tag{6.6}$$

where α is the spectral index of the power law, f_0 is some reference frequency and $\Omega_0 = \Omega(f = f_0)$. In particular stochastic backgrounds and foregrounds from compact binaries are expected to follow a power law with $\alpha = 2/3$ [264]. For the rest of this chapter, the power-law spectral shape is assumed to hold.

Spherical harmonics provide a general orthonormal basis to parameterize an arbitrary continuous and differentiable function on the two sphere:

$$\mathcal{P}(\hat{\mathbf{n}}) = \frac{1}{\sqrt{4\pi a_{0,0}}} \sum_{\ell,m} a_{\ell,m} Y_{\ell,m}(\hat{\mathbf{n}}), \tag{6.7}$$

where $\{Y_{\ell,m}\}$ are the spherical harmonic functions and the normalization factor of $\sqrt{4\pi a_{0,0}}$ ensures that the expansion satisfies Eq. 6.5. The coefficients $\{a_{\ell,m}\}$ are in general complex numbers and characterize the distribution of the field on the sky. The harmonics of positive and negative m are related by:

$$Y_{\ell,-m}(\hat{\mathbf{n}}) = (-1)^m Y^*_{\ell,m}(\hat{\mathbf{n}}). \tag{6.8}$$

The spherical harmonic functions form an orthonormal basis on the two sphere,

$$\int d^2n\, Y_{\ell,m}(\hat{\mathbf{n}}) Y^*_{\ell',m'}(\hat{\mathbf{n}}) = \delta_{\ell,\ell'}\delta_{m,m'}. \tag{6.9}$$

6.2.1 Detector response in the spherical harmonic basis

Since GW power is decomposed in the orthonormal spherical harmonic basis, it is also useful to calculate the detector response to each spherical harmonic mode. The expressions derived in Sec. 1.4.1 will be reused. The antenna pattern function of channel I of the detector is defined as [30, 265],

$$F_I^A(f,t,\hat{\mathbf{n}}) = D_I^{\mathrm{ij}}(f,t,\hat{\mathbf{n}}) : e_{\mathrm{ij}}^A(\hat{\mathbf{n}}), \tag{6.10}$$

where $D_I^{\mathrm{ij}}(f,t,\hat{\mathbf{n}})$ is the detector response tensor. For Michelson channels with arm orientations given by unit vectors $\hat{\mathbf{u}}$ and $\hat{\mathbf{v}}$, following Eq. 1.31, the response tensor is given by,

$$D_{\mathrm{mich}}(f,t,\hat{\mathbf{n}}) = \frac{1}{2}\Big[(\hat{\mathbf{u}}\otimes\hat{\mathbf{u}})\,\mathcal{T}(f,\hat{\mathbf{u}}\cdot\hat{\mathbf{n}}) - (\hat{\mathbf{v}}\otimes\hat{\mathbf{v}})\,\mathcal{T}(f,\hat{\mathbf{v}}\cdot\hat{\mathbf{n}})\Big], \tag{6.11}$$

assuming that the satellite motion is negligible during the round-trip light-travel time between the satellites. The response tensor is a function of time owing to the temporal variation of $\hat{\mathbf{u}}$ and $\hat{\mathbf{v}}$ as the satellites move in their orbits. Here, $\mathcal{T}(f,\hat{\mathbf{u}}\cdot\hat{\mathbf{n}})$ is the timing transfer function of interferometric detectors to GWs, which for an equal arm detector is given by Eq. 1.29 reproduced below again [266, 30] :

$$\begin{aligned}\mathcal{T}(f,\hat{\mathbf{u}}\cdot\hat{\mathbf{n}}) = \frac{1}{2}\Bigg[&\mathrm{sinc}\left(\frac{f}{f_*}(1-\hat{\mathbf{n}}\cdot\hat{\mathbf{u}})\right)\exp\left(-i\frac{f}{f_*}(3+\hat{\mathbf{n}}\cdot\hat{\mathbf{u}})\right)\\ &+\mathrm{sinc}\left(\frac{f}{f_*}(1+\hat{\mathbf{n}}\cdot\hat{\mathbf{u}})\right)\exp\left(-i\frac{f}{f_*}(1+\hat{\mathbf{n}}\cdot\hat{\mathbf{u}})\right)\Bigg],\end{aligned} \tag{6.12}$$

where $L = 2.5\times10^9$ m is the arm length of LISA and $f_* \equiv c/\pi L$. Since the SGCN is measured as the excess GW power in a detector, to detect it, we will need to correlate two channels I and J [3]. The response function to power distributed as $\mathcal{P}(\hat{\mathbf{n}})$ on the sky is then given by,

[3] Note that I and J can be the same channel of LISA. In the case of LIGO and Virgo, we usually only correlate distinct spatially-seperated interferometers, so the correlation response function is called the overlap reduction function.

$$R^{IJ}(f,t) = \int \frac{d^2n}{4\pi}\,\mathcal{P}(\hat{\mathbf{n}})\left(\sum_A F_I^A(f,\hat{\mathbf{n}})F_J^{A*}(f,\hat{\mathbf{n}})\right). \tag{6.13}$$

Expanding $\mathcal{P}(\hat{\mathbf{n}})$ in the spherical harmonic basis as in Eq. 6.7 we can define the response function to the $Y_{\ell,m}$ mode to be,

$$\mathcal{R}^{IJ}_{\ell,m}(f,t) = \int \frac{d^2n}{4\pi}\,Y_{\ell,m}(\hat{\mathbf{n}})\left(\sum_A F_I^A(f,\hat{\mathbf{n}})F_J^{A*}(f,\hat{\mathbf{n}})\right). \tag{6.14}$$

Finally combining Eqs. 6.3, 6.4, 6.7 and 6.14, the SGCN signal in the correlation between channels I and J is,

$$S^{\rm GW}_{IJ}(f,t) = \frac{3H_0^2}{2\pi^2 f^3}\frac{\Omega(f)}{\sqrt{4\pi a_{0,0}}}\sum_{\ell,m} a_{\ell,m}\mathcal{R}^{IJ}_{\ell,m}(f,t). \tag{6.15}$$

6.3 Clebsch-Gordan decomposition

The general spherical harmonic expansion describes a complex field. We can constrain it to be real everywhere on the sky with the condition:

$$a_{\ell,-m} = (-1)^m a^*_{\ell,m}. \tag{6.16}$$

However, the decomposition of GW power needs to be not only real but also non-negative for any direction on the sky, i.e., $\Omega(f,\hat{\mathbf{n}}) \geq 0$. This is especially important for LISA, as the GW power distribution will be highly anisotropic due to the foreground from galactic binaries. Implementing this constraint is also necessary for Bayesian inference with the spherical harmonic basis. This is because the posterior should be zero for any set of $a_{\ell,m}$'s that contains even the tiniest spot on the sky with negative GW power. Previous work in the PTA band attempted to solve this problem by numerically checking the sign of the GW power on the sky using a grid [267, 268] and assigning a probability of zero for a given set of $a_{\ell,m}$'s if any of the pixels have negative power. However, one can always use finer and finer grids to check this which makes this solution computationally untenable. A solution that mathematically guarantees non-negative power is much preferable. This section describes a solution to this problem using Clebsch–Gordan coefficients.

First, let us define a function $\mathcal{S}(\hat{\mathbf{n}})$ which is the square root of the spherical harmonic

expansion, i.e,

$$\mathcal{S}(\hat{\mathbf{n}}) = \left[\sum_{\ell,m} a_{\ell,m} Y_{\ell,m}(\hat{\mathbf{n}})\right]^{1/2}. \tag{6.17}$$

We then expand $\mathcal{S}(\hat{\mathbf{n}})$ through its own spherical harmonic expansion as,

$$\mathcal{S}(\hat{\mathbf{n}}) = \sum_{\ell,m} b_{\ell,m} Y_{\ell,m}(\hat{\mathbf{n}}). \tag{6.18}$$

The necessary and sufficient condition for the GW power to be non-negative then becomes $\mathcal{S}(\hat{\mathbf{n}}) \in \mathbb{R}$ which implies that $b_{\ell,-m} = (-1)^m b^*_{\ell,m}$. From Eqs. 6.17 and 6.18 we get,

$$\sum_{L,M} a_{L,M} Y_{L,M} = \left(\sum_{\ell,m} b_{\ell,m} Y_{\ell,m}(\hat{\mathbf{n}})\right)^2. \tag{6.19}$$

Expanding the right-hand side gives us:

$$\sum_{L,M} a_{L,M} Y_{L,M} = \sum_{\ell,m}\sum_{\ell',m'} b_{\ell,m} b_{\ell',m'} Y_{\ell,m}(\hat{\mathbf{n}}) Y_{\ell',m'}(\hat{\mathbf{n}}). \tag{6.20}$$

The Clebsch-Gordan coefficients $C^{LM}_{\ell m,\ell' m'}$ provide the means to write products of spherical harmonics as a sum over spherical harmonics [4], a trick that has been used in GW literature before but in different contexts [269, 270]. Here we get ;

$$Y_{\ell,m}(\hat{\mathbf{n}}) Y_{\ell',m'}(\hat{\mathbf{n}}) = \sum_{L=L_{\rm min}}^{L_{\rm max}} \sqrt{\frac{(2\ell+1)(2\ell'+1)}{4\pi(2L+1)}}\, C^{LM}_{\ell m,\ell' m'}\, C^{L0}_{\ell\, 0,\ell'\, 0} Y_{L,M}(\hat{\mathbf{n}}). \tag{6.21}$$

The expansion obeys selection rules that are related to the symmetries of the rotation group SO(3) [5]. These rules can be listed as:

[4] Alternatively one can use Wigner-3j symbols

[5] Ch. 16 of [179] contains a detailed discussion of the mathematics of spherical harmonics and the Clebsch-Gordan coefficients.

- $M = m + m'$
- $L_{\rm min} = \min(|\ell - \ell'|, |m + m'|)$ and $L_{\rm max} = \ell + \ell'$
- L is an integer

For compactness, let us define $\beta^{L,M}_{\ell m,\ell' m'}$ such that:

$$\beta^{L,M}_{\ell m,\ell' m'} = \sqrt{\frac{(2\ell+1)(2\ell'+1)}{4\pi(2L+1)}} C^{LM}_{\ell m,\ell' m'} C^{L0}_{\ell\, 0,\ell'\, 0}, \tag{6.22}$$

when the selection rules are satisfied, but $\beta^{L,M}_{\ell m,\ell' m'} = 0$ otherwise. We can then write Eq. 6.21 as,

$$Y_{\ell,m}(\hat{\mathbf{n}}) Y_{\ell',m'}(\hat{\mathbf{n}}) = \sum_{L,M} \beta^{L,M}_{\ell m,\ell' m'} Y_{L,M}(\hat{\mathbf{n}}). \tag{6.23}$$

Combining this with equation Eq. 6.20 gives:

$$\begin{aligned} \sum_{L,M} a_{L,M} Y_{L,M}(\hat{\mathbf{n}}) = \sum_{L,M} \left(\sum_{\ell m} \sum_{\ell' m'} b_{\ell,m} b_{\ell',m'} \beta^{\ell m,\ell' m'}_{L,M} \right) \\ \times Y_{L,M}(\hat{\mathbf{n}}). \end{aligned} \tag{6.24}$$

Since the set of Y_{LM} form an orthonormal basis, this provides the recipe for converting between $a_{\ell,m}$ and $b_{\ell,m}$:

$$a_{L,M} = \sum_{\ell,m} \sum_{\ell',m'} b_{\ell,m} b_{\ell',m'} \beta^{\ell m,\ell' m'}_{L,M}. \tag{6.25}$$

Often, we want to impose an artificial cutoff on angular sensitivity of some $\ell^a_{\rm max}$ on the expansion in $a_{\ell,m}$'s. This cutoff can correspond to estimated resolution limits of the detector itself, or it might be astrophysically motivated. The corresponding cutoff of the expansion on the $b_{\ell,m}$'s is taken to be $\ell^b_{\rm max} = \ell^a_{\rm max}/2$ assuming that $\ell^a_{\rm max}$ is an even number. This is a consequence of the second selection rule that $L_{\rm max} = \ell + \ell'$, which implies that the cutoff on the expansion in the $b_{\ell,m}$'s should be half that of the $a_{\ell,m}$'s if we want all terms higher than $\ell^a_{\rm max}$ in the latter to be zero. For the rest of this chapter we will use this relation and assume $\ell^a_{\rm max}$ is even.

Since $\Omega(f, \hat{\mathbf{n}})$ is proportional to $\mathcal{S}^2(\hat{\mathbf{n}})$, it is invariant under any transformation which leaves the latter invariant. The constraint that $\mathcal{S}(\hat{\mathbf{n}})$ be real [6] makes $\Omega(f, \hat{\mathbf{n}})$ invariant under a parity transformation $\{b_{\ell,m}\} \to \{-b_{\ell,m}\}$. However, this leftover symmetry introduces degeneracies that induce multiple modes in the posterior distribution. Moreover, since the $\{a_{\ell,m}\}$ expansion is normalized as in Eq. 6.5, there is also a scale invariance within the $\{b_{\ell,m}\}$ space, i.e. with the transformation $\{b_{\ell,m}\} \to \{\kappa\, b_{\ell,m}\}$, where κ is some constant. One can break both these symmetries by fixing the value of one of the $b_{\ell,m}$ coefficients. In this chapter we choose to fix $b_{0,0} = 1$.

A similar method to what is used here can also be applied to expand a probability distribution or any other non-negative function on the two sphere. Most of the results in this section will generalize to that case with the additional requirement that the normalization of the distributions be one. Indeed, the Clebsch-Gordan based spherical harmonic decomposition was recently used to constrain anisotropies in the distribution of BBH progenitors using events detected by LIGO-Virgo up-to-the second observing run [192], and also to develop an optimized anisotropic pipeline for PTAs [263].

6.4 BLIP pipeline

This section briefly introduces the BLIP pipeline [7], which is an independent Python-based implementation for LISA data analysis and the details of it. BLIP is designed to be a general-purpose LISA data-analysis pipeline but is used specifically for SGCN analysis in this chapter. It is written to make it easy to add new GW signal models and likelihood models along with simulating instrumental Gaussian noise in the time domain. The instrumental noise is simulated as the sum of acceleration and position noises using the spectral form described in the LISA proposal [231]. The functional forms for the power spectrum of the acceleration and position noise are given by

[6] In fact we could have chosen $\mathcal{S}$ such that $\Omega = |\mathcal{S}|^2$ in which case $\mathcal{S}$ could be any general complex number. The degree of degeneracies would be higher, so it would be a less economical choice than what we actually make.

[7] https://github.com/sharanbngr/blip

$$
\begin{aligned}
S_p(f) =& N_p \left[1 + \left(\frac{2\,\mathrm{mHz}}{f}\right)^4\right] \mathrm{Hz}^{-1}, \\
S_a(f) =& \left[1 + \left(\frac{0.4\,\mathrm{mHz}}{f}\right)^2\right] \left[1 + \left(\frac{f}{8\,\mathrm{mHz}}\right)^4\right] \\
& \times \frac{N_a}{(2\pi f)^4} \mathrm{Hz}^{-1}
\end{aligned}
\tag{6.26}
$$

The noise levels can be set by the end-user by modifying N_p and N_a in the pipeline, but the current implementation assumes that they are the same in all satellite links. In this chapter we set $N_p = 9 \times 10^{-42}$ and $N_a = 3.6 \times 10^{-49}$ Hz^{-4} to match the instrumental noise levels described in the LISA proposal. The code also implements time-delay interferometry (TDI) with Michelson, $X - Y - Z$ and $A - E - T$ [271, 272] channels and heliocentric rigid-body orbits of the LISA satellites. This is implemented in an adiabatic manner by modeling the satellites to be stationary for small segments of time ($\ll$ 1 year) both for signal simulation and recovery, and allowing them to move between the time segments. The current implementation of orbits neglects the differential time delay for laser light on the round trip between two satellites – i.e the travel time difference from satellite A to B and from B to A – and also neglects their breathing modes. The satellites' orbital motion is especially important in partially breaking the degeneracies of the antenna patterns when detecting an anisotropic SGCN.

The strain data $d_I(t)$ from any channel I is the sum of the instrumental noise $n_I(t)$ and the GW signal in the channel $h_I^{GW}(t)$,

$$
d_I(t) = n_I(t) + h_I^{GW}(t). \tag{6.27}
$$

Due to the linear nature of the Fourier transform this relationship carries over to the Fourier domain as well,

$$
\tilde{d}_I(f) = \tilde{n}_I(f) + \tilde{h}_I^{GW}(f). \tag{6.28}
$$

In the case of SGCNs, due to the Gaussian nature of both the GW strain and the instrumental noise and under the assumption that they are uncorrelated, the PSD of the data is the sum of the data and the noise PSDs,

$$S_{II}(f) = S^{n}_{II}(f) + S^{GW}_{II}(f). \tag{6.29}$$

A similar relation will also hold for cross correlation between channels I and J too,

$$S_{IJ}(f) = S^{n}_{IJ}(f) + S^{GW}_{IJ}(f). \tag{6.30}$$

Finally, the BLIP pipeline is built to facilitate Bayesian inference and supports emcee [273] and dynesty [110] samplers. The pipeline is for the most part sampler agnostic, making it easy to add support for additional statistical samplers. All results in this chapter were made through the dynesty sampler. The Clebsch–Gordan coefficients are implemented with the help of the Wigner module of SymPy [274].

6.4.1 Analysis configuration

For the remainder of this chapter we will use $X-Y-Z$ TDI channels with an equal-arm rigid-body orbiting configuration of LISA. For such a configuration, the X channel is related to the Michelson channel at vertex 1 as,

$$X(t) = M_1(t) - M_1(t - 2L/c). \tag{6.31}$$

This can be written in the frequency domain as,

$$\tilde{X}(f) = -2i\tilde{M}_1(f)\sin\left(\frac{\pi f}{f_*}\right)\exp\left(-\frac{i\pi f}{f_*}\right). \tag{6.32}$$

The channels Y and Z are similarly related to the Michelson channels M_2 and M_3.

To analyze the simulated data, we employ Fourier transforms with a duration of $T_{\rm seg} = 10^5$ s with the aforementioned adiabatic approximation within each segment. The sampling frequency of the data is $f_s = 0.25$ Hz. We only consider the Fourier components in between $f_{\rm min} = 2\times10^{-4}$ Hz and $f_{\rm max} = 2\times10^{-2}$ Hz for the analysis. We also want to approximate the covariance matrix of the data in this frequency-time analysis to be diagonal across frequency and time, for which we require the auto-correlation time scale to be much smaller than $T_{\rm seg}$ [11]. The value of $T_{\rm seg} = 10^5$ s is thus chosen as a compromise between the auto-correlation time scale of the noise, which is $\sim 10^4$ s for this frequency band, and the adiabatic approximation of the motion of the satellites.

Each time segment is Hann-windowed before Fourier transformation.

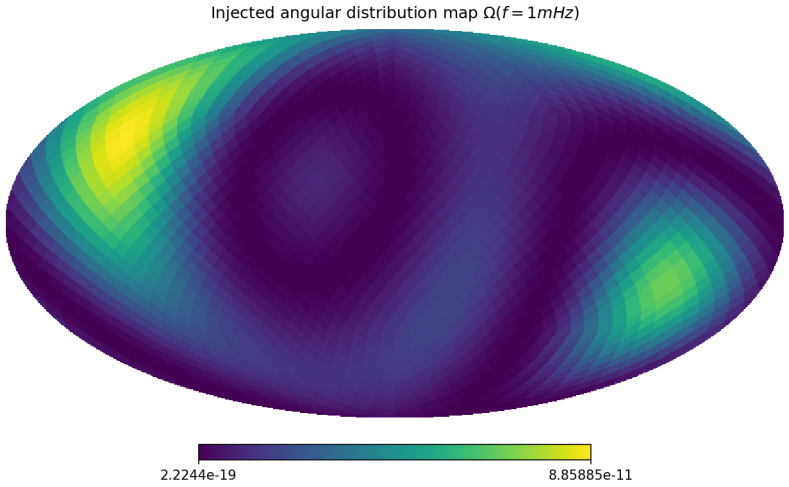
6.4.2 Likelihood function

The power spectral density (PSD) and the cross-spectral density (CSD) of the data are combinations of the GW power from the SGCN and of the instrumental noise power [98]. Assuming that both are Gaussian, the Fourier domain likelihood is based on the multi-dimensional complex Gaussian distribution as described by Eq. 2.17:

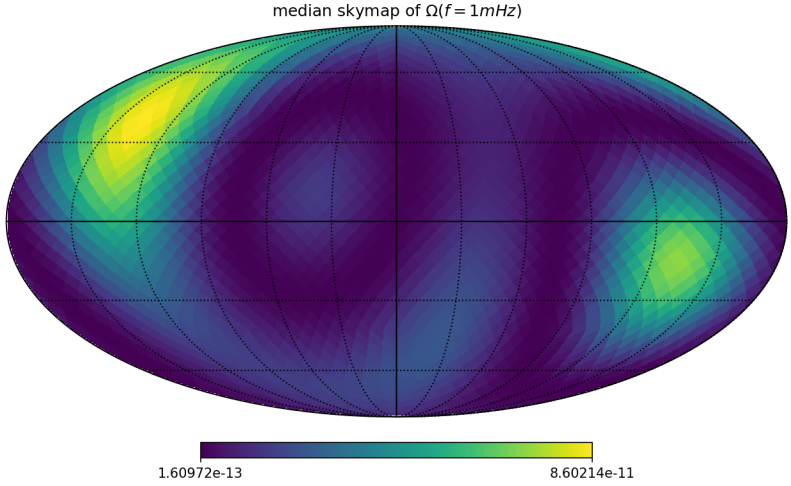
$$\mathcal{L}(\tilde{d}|N_p, N_a, \{b_{\ell,m}\}) = \prod_{t,f} \frac{1}{2\pi T_{\rm seg}^3 |C(t,f)|} \times \exp\left(-\frac{2\,\tilde{d}^*_{t,f}\, C(t,f)^{-1}\, \tilde{d}_{t,f}}{T_{\rm seg}}\right). \tag{6.33}$$

Here $\tilde{d}_{t,f} = [\tilde{d}_X(t,f), \tilde{d}_Y(t,f), \tilde{d}_Z(t,f)]$ is the array of data in the Fourier domain for the three channels measured in the time segment labeled by t and at frequency f. As previously mentioned, the data is Fourier transformed in segments of duration $T_{\rm seg}$, and the product is across all frequency bins and time-segments. The term $C(t,f)$ is the 3×3 covariance matrix across the three channels, and as seen in Eqs. 6.29 and 6.30, its elements are the sum of the signal spectral densities (defined in Eq. 6.15) and the instrumental noise spectral densities $S^n_{IJ}(f)$. We follow the derivation in [272] of the actual expressions for the noise spectral densities.

The term $|C(t,f)|$ is the determinant of the covariance matrix. In many applications of Bayesian inference to GW data, one usually can ignore the overall normalization because it is constant and model independent. This is not true in this case and correctly modeling the normalization is essential to the problem of Bayesian inference of the SGCN.



(a)



(b)

Figure 6.1: Skymaps for the simulation and analysis described in Sec. 6.5 with the analysis time scale of one year. The map on the top is the simulated skymap while the bottom shows the posterior median recovery skymap. Both maps show the distribution of $\Omega(f = 1 \text{ mHz})$ in the solar system barycentric frame. The full posteriors corresponding to these maps are shown in Fig. 6.3. Figure reproduced from Ref. [13] with permission.

6.5 Simulation and detection

6.5.1 Validation

We first validate the Clebsch–Gordan technique by recovering an ad hoc distribution of power, simulated with a power-law spectral-index SGCN with $\ell^a_{\rm max} = 4$. The simulated $\{b_{\ell,m}\}$ coefficients are [8] (1.0, 0.75, 0.5, $0.7j$, $0.7 - 0.3j$, $1.1j$) which yield the GW power distribution on the sky shown in Fig. 6.1. These values are chosen only to validate the ability of the algorithm to recover an arbitrary distribution of power on the sky. The spectral index of the power law is chosen to be consistent with binary inspiral at $\alpha = 2/3$ with $\Omega(f = 25\text{Hz}) = 2 \times 10^{-7}$.

The spherical harmonic coefficients $\{b_{\ell,m}\}$ are complex if $m \neq 0$ and thus have two degrees of freedom. We parameterize them by their amplitude $|b_{\ell,m}|$ and phase $\phi_{\ell,m}$. We set uniform priors between $[0, 3]$ on the amplitude and uniform priors between $[-\pi, \pi]$ on the phase. For the modes with $m = 0$, i.e $b_{\ell,0}$'s which are real, we set uniform priors between $[-3, 3]$. The variance of the prior sky map due to these choices of priors is

[8] in healpix order

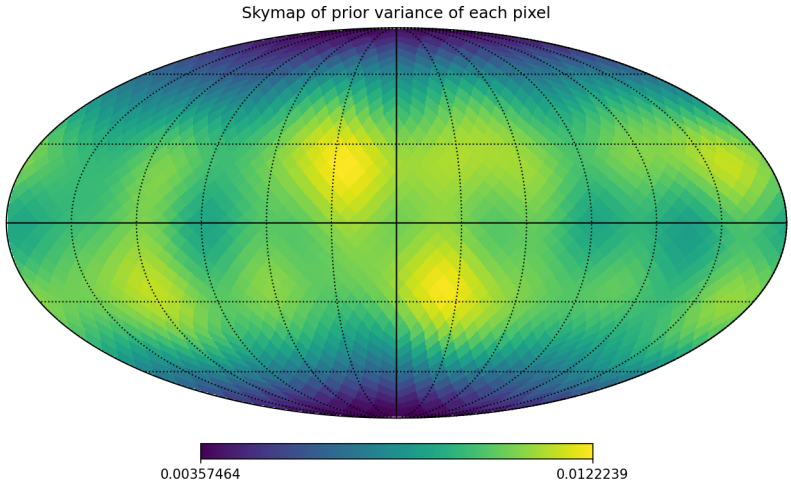


Figure 6.2: The variance of $a_{\ell,m}$'s sampled from the prior distribution is shown here, without the multiplicative factor of $\Omega(f)$. The nearly isotropic distribution of the variance across the sky demonstrates that the priors have broad support for multiple spherical harmonic modes. Figure reproduced from Ref. [13] with permission.

shown in Fig. 6.2. This shows broad support for many modes and implies that the priors are not peaked at any particular region in the space.

We use the single channel signal-to-noise ratio (SNR) as a metric to characterize the strength of the signal, defined as

$$\mathrm{SNR} = \sqrt{T_{\mathrm{seg}} \sum_{t,f} \delta f \, \frac{S^{GW}_{XX}}{S^{n}_{XX}}} \tag{6.34}$$

where the summation is over time and frequency band used for analyzing the data. Note that while the SNR is a metric of the strength of the SGCN, because of the non-isotropic distribution of the SGCN, it does not grow as $\sim T^{1/2}$ as would be expected in the isotropic case.

Two separate analyses with a duration of one year and two months were ran. Figures 6.3 and 6.4 show the posterior corner plots for the former and the latter respectively. Bimodalities are observed in the posteriors for the $b_{\ell,m}$ in the two-month run, which are related to the parity symmetry described in Sec. 6.3. While the choice of fixing $b_{0,0} = 1$ breaks this symmetry in principle, in practice, this breaking is only approximate and can fail in the limit of a weak signal or a small amount of observation time. This is because while two modes may be related by a sign change due to parity, we are limited by the ability of LISA to resolve this relative sign (with respect to $b_{0,0}$). However, with enough time or a stronger signal, the breaking of the symmetry becomes complete, and the sampler finds the right mode as seen in the posterior for the one year run. The corresponding injected skymap, and the recovered median posterior skymap (for 1 year long run) are shown in Fig. 6.1. The median posterior map is the skymap generated from median values of posterior samples.

6.5.2 Localized sources

Next, we test the ability to recover signals from localized sources of GW power, often referred to as point sources in the literature. Examples of such sources could be combined gravitational radiation from a large globular cluster or a nearby galaxy. The GW power from these sources would be an incoherent superposition of all the GWs emitted by individual sources within them. Such sources are generally localized to much smaller

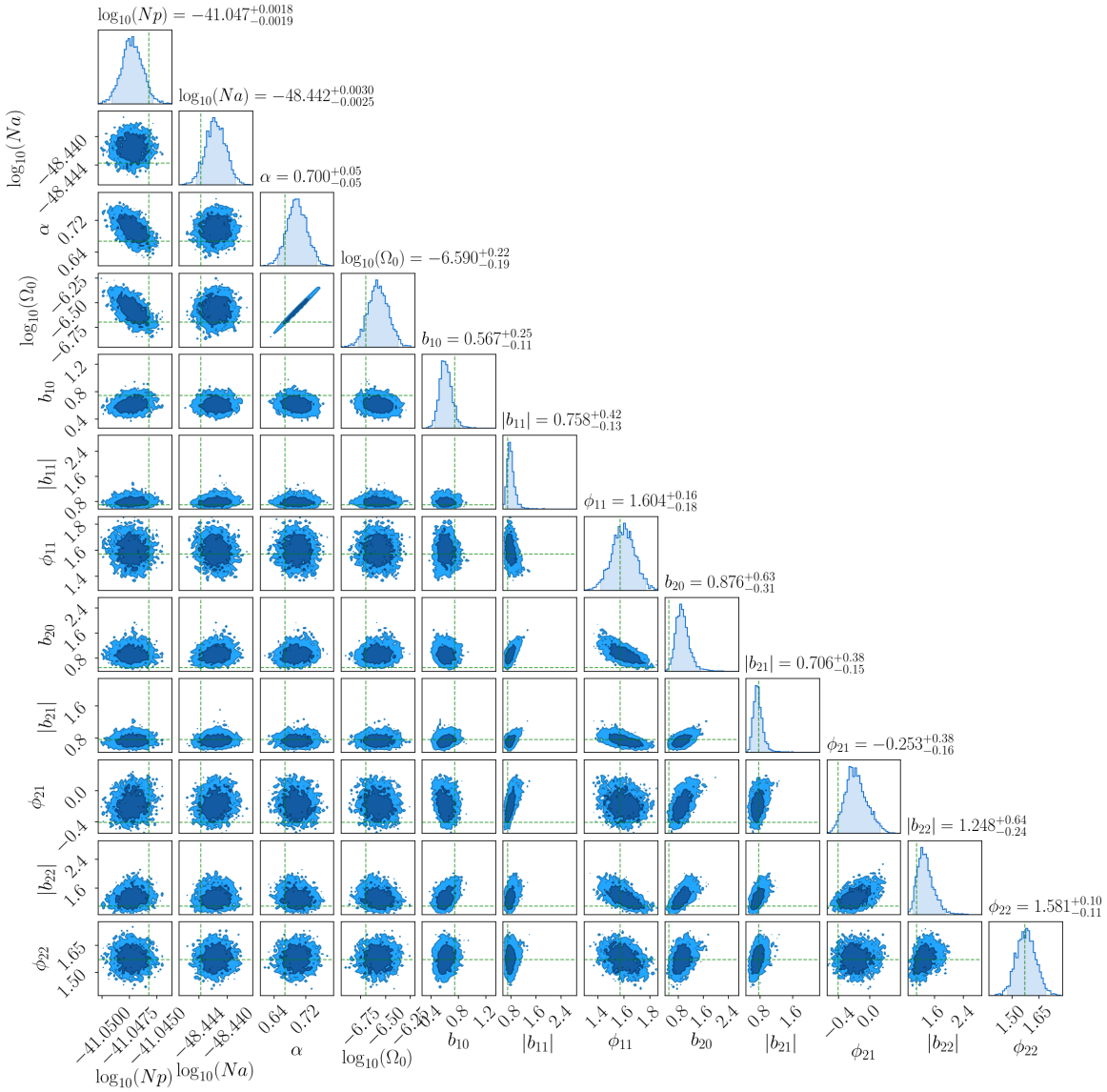


Figure 6.3: Posteriors corresponding to skymaps in Fig. 6.1 with a signal amplitude $\Omega_0 = 2\times10^{-7}$ for the duration of 1 year. This corresponds to a single channel theoretical SNR ≈ 149. The shaded region in the one-dimensional posteriors are 95% confidence levels while the light and dark regions in the two-dimensional posteriors are one and two sigma confidence levels respectively. The parameters N_p and N_a are the posterior measurements for the position and acceleration noises respectively using the functional forms described in Eq. 6.26. The parameters α and Ω_0 measure the spectral shape of the SGCN while the rest of the parameters are measurements of $b_{\ell,m}$'s which describe the distribution of GW power on the sky. The dashed green lines are the true values of these parameters used when simulating the data. Figure reproduced from Ref. [13] with permission.

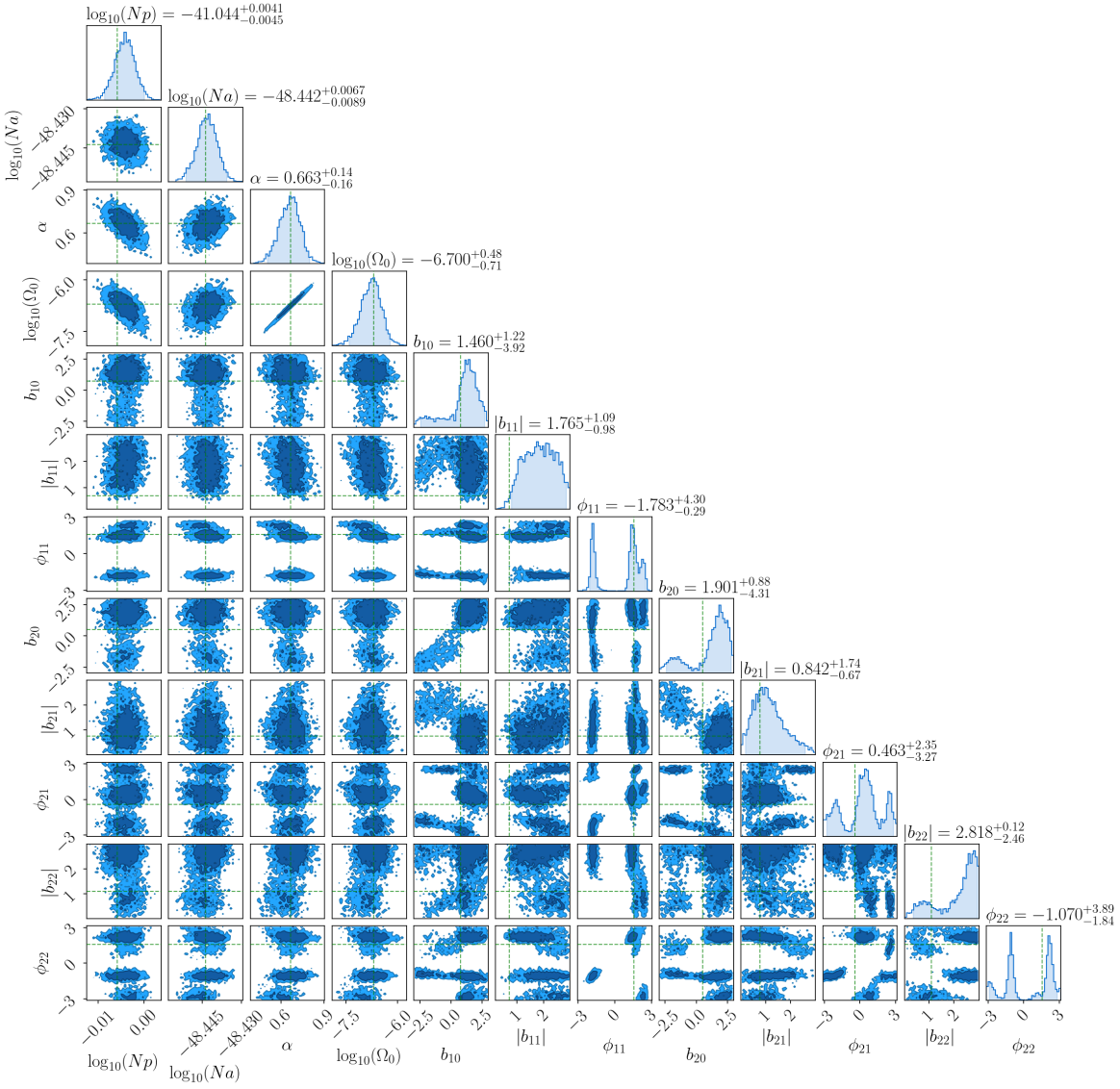


Figure 6.4: Posteriors for an analysis with a signal amplitude $\Omega_0 = 2 \times 10^{-7}$ for the duration of 2 months. This corresponds to a single channel theoretical SNR ≈ 59. The shaded region in the one-dimensional posteriors are 95% confidence levels while the light and dark regions in the two-dimensional posteriors are one and two sigma confidence levels respectively. The bimodalities seen in the one-dimensional posteriors are due to the parity symmetries described in Sec. 6.3 which are only approximately broken in the limit of a weak signal or short integration time. With a stronger signal, the breaking of the symmetry becomes more complete and the degenerate modes go away as the sampler finds the right mode, as can be seen in the posterior for the 1 year run (Fig. 6.3). Figure reproduced from Ref. [13] with permission.

angular scales than is possible to resolve with LISA. Thus, when they are mapped with spherical harmonic methods, the power is usually smeared on larger angular scale as determined by the strength of the signal, the integration time and the $\ell^a_{\rm max}$ scale where we cutoff the expansion. The dependence on the latter is demonstrated in Fig. 6.5 where three months of simulated data containing a point source signal is analyzed using $\ell^a_{\rm max} = 4$ and 6, with the same $\Omega(f)$ in both cases. As expected the smearing of the power is smaller in the case of $\ell^a_{\rm max} = 6$. In both cases, we still assumed $\alpha = 2/3$ power-law spectrum.

6.6 Galactic Foreground

A significant source of GW confusion noise in the LISA band is expected to be due to the foreground from the galactic DWD binaries. The foreground will peak above the instrumental noise floor at ~ 1 mHz, will be strongly anisotropic, and will exhibit temporal modulations as LISA orbits the sun. This section tests the Bayesian analysis developed thus far on a simplified model of the galactic foreground.

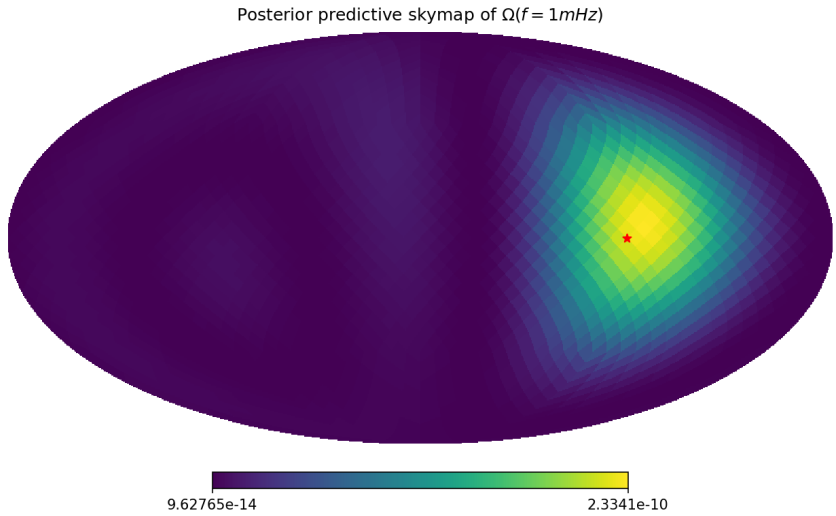
The spectral content of the foreground was simulated using an $\alpha = 2/3$ power law. This is a reasonable approximation at low frequencies where the inspiral dynamics are characterized by GW emission. However, it breaks down at frequencies of several mHz because of mass transfer with DWDs. The anisotropy of the foreground was simulated using the galactic distribution model developed in [275, 255, 276]. The distribution of the DWD sources in the galaxy is modeled as a disk with a central bulge. The density of the disk is given by,

$$\rho_d(r, z) \propto \exp\left(-r/r_h\right) \exp\left(-z/z_h\right). \tag{6.35}$$

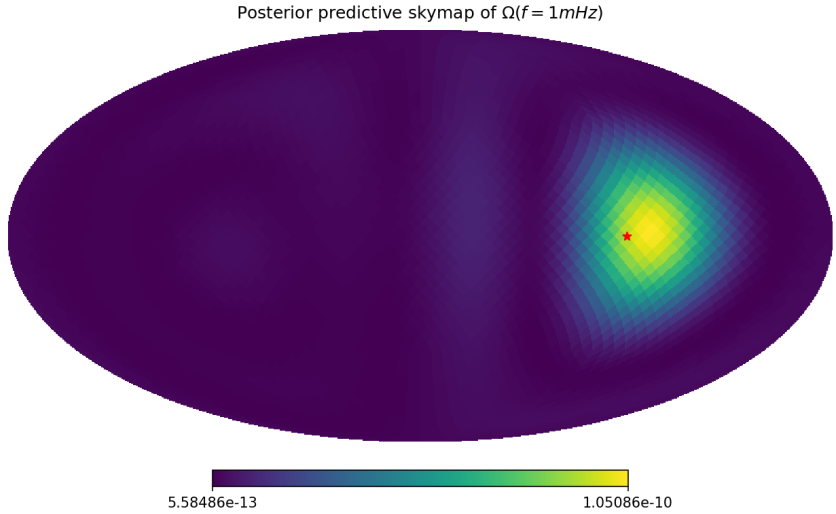
Here r_h and z_h are the radial and vertical scale height parameters assumed to be 2.9 kpc and 0.3 kpc, respectively, following the thin disk model of [275]. The galactic bulge is modeled to have an azimuthally symmetry,

$$\rho_b(r) \propto \frac{\exp\left(-(r/r_{\rm cut})^2\right)}{(1 + r'/r_0)^\gamma}, \tag{6.36}$$

where $\gamma = 1.8$, $r_0 = 0.075$ kpc, $r_{\rm cut} = 2.1$ kpc, $r' = \sqrt{r^2 + (z/q)^2}$ and $q = 0.5$.

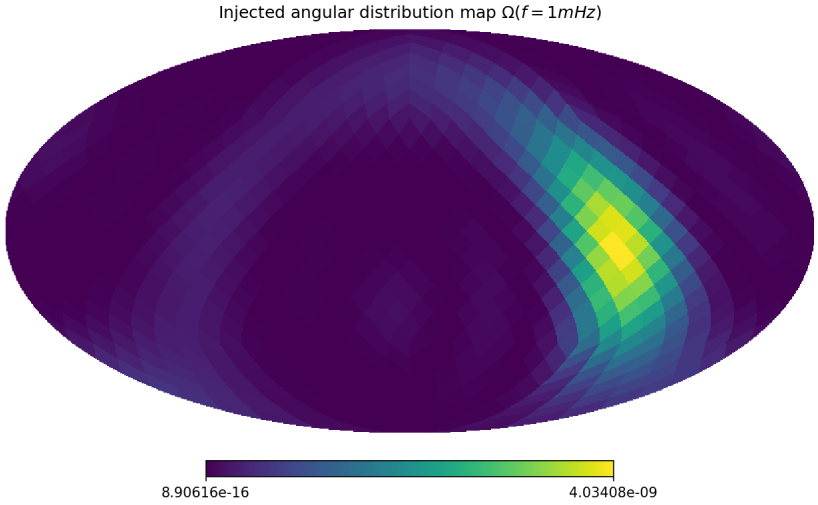


(a)

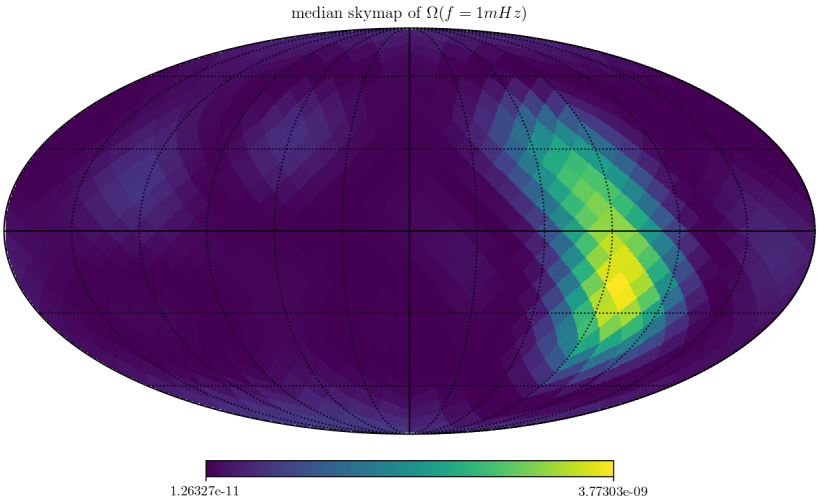


(b)

Figure 6.5: Recovered skymaps for simulated localized sources using three months of simulated data. The red star indicates the true position of the source. The map on the top uses a cutoff of $\ell^a_{max} = 4$ while the bottom uses $\ell^a_{max} = 6$. Consequently, the power is smeared to a larger extent on the sky for the former compared to the latter. Figure reproduced from Ref. [13] with permission.



(a)



(b)

Figure 6.6: Skymaps for the simulation and recovery of the galactic DWD foregrounds described in Sec. 6.6 with one year of data. The top panel shows the simulated skymap while the bottom shows posterior median recovery skymap. Both maps show the distribution of $\Omega(f = 1 \times 10^{-3}$ Hz) in the solar system barycentric frame. The bright spots in the map corresponds to the galactic central bulge. Figure generated with the help of Alexander Criswell and reproduced from Ref. [13] with permission.

Ideally, one would draw many independent DWD binaries from this model (and with the given spectral shape). However, in this simplified analysis, the galaxy was divided in a cartesian grid of size $200 \times 200 \times 200$, and the GW power in each cube was assumed to be proportional to the density of the model. The GW power is then multiplied by the inverse square of the distance to the Solar System and summed up over the grid to yield the anisotropic power distribution seen in the top panel of Fig. 6.6.

A year's worth of this data was then analyzed using the Bayesian spherical harmonic method. An $f_{\rm max}$ of 1 mHz was used to artificially approximate the mass transfer effect. The bottom panel of Fig. 6.6 shows the recovered sky map when using $\ell_{\rm max} = 4$. It is to be noted that the central bulge, in particular, is especially well resolved.

6.7 Discussion & Conclusion

This chapter developed a Bayesian mapping algorithm using the spherical harmonic basis that can optimally recover any arbitrary distribution of GW power on the sky using LISA data, while imposing the physical constraint that the GW power is non-negative in all directions on the sky. This method was validated through a series of end-to-end simulations of different types of SGCNs.

While this chapter primarily focused on developing general mapping tools, there are several directions to push this forward and apply them for astrophysical use. The galactic foreground simulated in this chapter is ultimately simplistic in that it assumes that it is Gaussian and has the spectral shape of a power law with $\alpha = 2/3$. It will be essential to relax these assumptions and validate this method on a realistic simulated foreground formed from catalogs of DWDs generated with population synthesis codes. This might also require relaxing the assumptions of Gaussianity. Additionally, as Ref.[255] shows, mapping the foreground on the sky can help constrain galactic structure in a manner complementary to resolvable DWDs while also probing the stellar evolution history of the galactic white dwarfs. It would be interesting to study what mapping the foreground could tell us about the astrophysics of galactic foreground with the optimal analysis developed in this work and using realistic simulations. One can also expand this to study the astrophysics of the galactic DWD population in a more model-independent manner by constraining the properties of the foreground in narrow frequency bands.

It will also be essential to study how the algorithm's angular sensitivity, characterized by the ℓ^a_{max} parameter, will scale with the strength of the stochastic confusion noise and the duration of the analysis. An optimal way to do this could be by making ℓ^a_{max} an independent parameter to be determined by data [9]. Since the dimensionality of the spherical harmonic parameter space itself depends on ℓ^a_{max}, such an analysis could use Reversible-Jump Markov-chain Monte Carlo approaches such as those developed in Ref. [99]. One can also imagine multiple SGCNs, arising from different sources that might be separable if they have different spectral shape and angular structure (for e.g. [277]).

Finally, an important application would be a joint analysis with both resolvable signals and the galactic DWD foreground. One of the core strengths of mapping the foreground in a Bayesian manner is the possibility of simultaneous inference of resolvable signals along with the galactic foreground. This will allow an unbiased estimation of the properties of the resolvable signals while accounting for the temporal modulation of the noise due to the galactic foreground.

[9] Note that this only determines the angular scale to which the noise can be confidently mapped on the sky in relation to the sensitivity of the detectors. It will not be a measure of a true angular cutoff of the stochastic noise, were that to be finite.

The Answer to the Great Question of Life, the Universe and Everything is forty-two.

Deep Thought, (Douglas Adams) The Hitchhiker's Guide to the Galaxy

Chapter 7

Conclusion and Discussion

This chapter summarizes the main ideas and results presented within this thesis and discusses potential extensions. We then finish with an overview of how these results fit into the context of GW astrophysics as a whole and the new science they could promise in the coming decades with both ground-based and space-based detectors.

Chapter 3 described an analysis to search for long-lived post-merger remnants. The BNS merger GW170817 gave us the first real shot at looking for a merger remnant through GWs but showed that the ground-based detectors were at least an order of magnitude off in strain sensitivity to start meaningfully constraining post-merger GW signals. This provides a key astrophysical motivation for next-generation detectors to achieve better sensitivity in the kHz band.

Once a GW transient is detected, we need parameter estimation to study its properties and understand what it can tell us about the astrophysics of its source. Chapter 4 developed Bayesian techniques for parameter estimation of poorly modeling transients. Methods were developed to fold in the modeling uncertainties in the GW phase while also accounting for PSD uncertainties that might be important when dealing with long-lived transients. However, the method used was phase incoherent, i.e., the phase between the detectors is not treated consistently when marginalized. An obvious improvement would be to incorporate the phase consistently, depending on the sky position, into the marginalizing. Appendix C derives a phase-coherent expression for what this will look like for a network of ground-based detectors.

Chapter 5 explores methods to use GWs for measuring the statistical properties of

the distribution of BBH mergers in the Universe. A two-point correlation method is developed, and its use is demonstrated to measure angular correlations of BBH mergers. It is also shown how this can be extended to cross-correlate BBH mergers with galaxy count tracers. The method developed is more general than just measuring angular correlations, and the most obvious extension is to measure 3D correlations, folding in GW distance measurements. The distance information will not just provide more accurate constraints but also a means to connect clustering measurements with physical models of structure formation.

In the 3D case, the BBH mergers act as tracers for the large-scale structure of the Universe, modulated by the redshift dependence of the merger rate. We can thus simultaneously infer the clustering parameters and the merger rate as a function of redshift, using, for example, star-formation rate-based models like the one developed by Madau & Dickinson [278, 279]. As GW detector sensitivity increases, and in particular with the advent of third-generation detectors, such "tomographic" GW measurements could be one of the best ways to constrain structure at $z > 1$ as galactic surveys do not extend to such distances [280].

Finally, Chapter 6 develops Bayesian methods for mapping the power on the GW sky with LISA. A novel decomposition is developed that uses Clebsch–Gordan coefficients to mathematically constrain the GW power to be non-negative everywhere in the spherical harmonic basis. This provides an optimal basis for mapping strongly anisotropic SGCNs, as demonstrated through a series of examples in the chapter. Several extensions and applications are possible for this work, especially when focusing on the galactic white dwarf foreground. Constraining the white dwarf distribution by measuring the stochastic gravitational waves foreground is one plausible application. As this is a Bayesian method, it also fits well with the analyses for resolvable signals. This makes joint analyses possible, allowing better accuracy and precision than either one of them can provide. Another interesting extension would be to attempt to constrain a cosmological stochastic background in the presence of the galactic foreground.

In the five years since the first direct detection of GWs, the nascent field of GW detection has already yielded several advances in astrophysics. The scientific scope of the GW astrophysics will only increase as sensitivity grows and as space-based detectors and the next generation of ground-based detectors are developed. The methods developed

here will help target key new science objectives for the future.

Some of the possible new science that can be probed from studies presented here can complement analyses that are already possible in interesting ways, with novel astrophysical or cosmological implications. For example, detecting the post-merger phase can allow us to study neutron star matter in a manner complementary to the BNS inspiral measurement of tidal Love numbers [25, 89]. Matter effects measured in the post-merger phase can break the degeneracy between inclination angle and distance measurement to a BNS merger, making a much more precise GW measurement of the Hubble constant possible [281].

The measurement of LSS, on the other hand, will add a new GW tool for cosmology, in addition to the current ability to measure the Hubble constant. The possibilities that open up by cross-correlation with galaxy count tracers are especially fascinating as the clustering properties can help separate different black hole populations [1]. An intriguing possibility could be to look for populations of black holes of primordial origin [282, 283] based on their clustering at high redshifts.

[1] This is separation at a statistical population-level. We will not be able to determine, of course, if any single BBH merger contains primordial progenitors.

References

[1] Robert E. Kass and Adrian E. Raftery. Bayes factors. *Journal of the American Statistical Association*, 90(430):773–795, 1995. ix, 34

[2] P. B. Covas et al. Identification and mitigation of narrow spectral artifacts that degrade searches for persistent gravitational waves in the first two observing runs of Advanced LIGO. *Phys. Rev. D*, 97(8):082002, 2018. ix, 26, 59, 62

[3] B. P. Abbott et al. Search for gravitational waves from a long-lived remnant of the binary neutron star merger GW170817. *Astrophys. J.*, 875(2):160, 2019. ix, xii, xiii, 46, 56, 57, 59, 64, 65, 68, 69, 70, 73, 74, 81, 162

[4] Aaron Buikema et al. Sensitivity and performance of the Advanced LIGO detectors in the third observing run. *Phys. Rev. D*, 102(6):062003, 2020. xi, 21, 22, 23, 25

[5] M. Tse et al. Quantum-Enhanced Advanced LIGO Detectors in the Era of Gravitational-Wave Astronomy. *Phys. Rev. Lett.*, 123(23):231107, 2019. xi, 24, 25

[6] D. Davis et al. LIGO Detector Characterization in the Second and Third Observing Runs. 1 2021. xi, 26, 27

[7] Michael W. Coughlin et al. Measurement and subtraction of Schumann resonances at gravitational-wave interferometers. *Phys. Rev. D*, 97(10):102007, 2018. xi, 27, 28

[8] B. P. Abbott et al. A gravitational-wave standard siren measurement of the Hubble constant. *Nature*, 551(7678):85–88, 2017. xi, 30, 105

[9] Tim Dietrich, Tanja Hinderer, and Anuradha Samajdar. Interpreting Binary Neutron Star Mergers: Describing the Binary Neutron Star Dynamics, Modelling Gravitational Waveforms, and Analyzing Detections. *Gen. Rel. Grav.*, 53(3):27, 2021. xi, 47, 48

[10] B. P. Abbott et al. Properties of the binary neutron star merger GW170817. *Phys. Rev. X*, 9(1):011001, 2019. xi, 49, 50, 69, 73

[11] Sharan Banagiri, Michael W. Coughlin, James Clark, Paul D. Lasky, M. A. Bizouard, Colm Talbot, Eric Thrane, and Vuk Mandic. Constraining the Gravitational-Wave Afterglow From a Binary Neutron Star Coalescence. *Mon. Not. Roy. Astron. Soc.*, 492(4):4945–4951, 2020. xiii, xiv, xvii, 35, 72, 75, 80, 81, 82, 118, 166, 167

[12] Sharan Banagiri, Vuk Mandic, Claudia Scarlata, and Kate Z. Yang. Measuring angular N-point correlations of binary black hole merger gravitational-wave events with hierarchical Bayesian inference. *Phys. Rev. D*, 102(6):063007, 2020. xiv, xv, 86, 99, 100, 102

[13] Sharan Banagiri, Alexander Criswell, Tommy Kuan, Vuk Mandic, Joseph D. Romano, and Stephen R. Taylor. Mapping the Gravitational-wave Sky with LISA: A Bayesian Spherical Harmonic Approach. 3 2021. xv, xvi, xvii, 106, 120, 121, 123, 124, 126, 127

[14] Albert Einstein. Die Feldgleichungen der Gravitation. *Sitzungsberichte der Königlich Preußischen Akademie der Wissenschaften (Berlin*, pages 844–847, January 1915. 1

[15] A. Einstein. Die Grundlage der allgemeinen Relativitätstheorie. *Annalen der Physik*, 354(7):769–822, January 1916. 2

[16] F. W. Dyson, A. S. Eddington, and C. Davidson. A Determination of the Deflection of Light by the Sun's Gravitational Field, from Observations Made at the Total Eclipse of May 29, 1919. *Philosophical Transactions of the Royal Society of London Series A*, 220:291–333, January 1920. 2

[17] R. V. Pound and J. L. Snider. Effect of gravity on nuclear resonance. *Phys. Rev. Lett.*, 13:539–540, Nov 1964. 2

[18] Irwin I. Shapiro, Gordon H. Pettengill, Michael E. Ash, Melvin L. Stone, William B. Smith, Richard P. Ingalls, and Richard A. Brockelman. Fourth test of general relativity: Preliminary results. *Phys. Rev. Lett.*, 20:1265–1269, May 1968. 2

[19] Albert Einstein. Näherungsweise Integration der Feldgleichungen der Gravitation. *Sitzungsberichte der Königlich Preußischen Akademie der Wissenschaften (Berlin*, pages 688–696, January 1916. 2

[20] Albert Einstein. Über Gravitationswellen. *Sitzungsberichte der Königlich Preußischen Akademie der Wissenschaften (Berlin*, pages 154–167, January 1918. 2

[21] J. H. Taylor, L. A. Fowler, and P. M. McCulloch. Measurements of general relativistic effects in the binary pulsar PSR1913 + 16. *nat*, 277(5696):437–440, February 1979. 3

[22] J. H. Taylor and J. M. Weisberg. A new test of general relativity Gravitational radiation and the binary pulsar PSR 1913+16. *apj*, 253:908–920, February 1982. 3

[23] B. P. Abbott et al. Observation of Gravitational Waves from a Binary Black Hole Merger. *Phys. Rev. Lett.*, 116(6):061102, 2016. 3, 73, 87

[24] J. Aasi et al. Advanced LIGO. *Class. Quant. Grav.*, 32:074001, 2015. 3, 21, 23, 73, 108

[25] B. P. Abbott et al. GW170817: Observation of Gravitational Waves from a Binary Neutron Star Inspiral. *Phys. Rev. Lett.*, 119(16):161101, 2017. 3, 26, 29, 30, 46, 47, 73, 79, 87, 132

[26] F. Acernese et al. Advanced Virgo: a second-generation interferometric gravitational wave detector. *Class. Quant. Grav.*, 32(2):024001, 2015. 3, 21, 73, 108

[27] B. P. Abbott et al. Multi-messenger Observations of a Binary Neutron Star Merger. *Astrophys. J. Lett.*, 848(2):L12, 2017. 3, 26, 29, 46, 47, 79

[28] M. Maggiore. *Gravitational Waves: Volume 1: Theory and Experiments*. 2008. 4, 6, 7, 20, 65, 110

[29] F. B. Estabrook and H. D. Wahlquist. Response of Doppler spacecraft tracking to gravitational radiation. *General Relativity and Gravitation*, 6(5):439–447, October 1975. 10

[30] Neil J. Cornish and Shane L. Larson. Space missions to detect the cosmic gravitational wave background. *Class. Quant. Grav.*, 18:3473–3496, 2001. 10, 112

[31] R. Schilling. Angular and frequency response of LISA. *Class. Quant. Grav.*, 14:1513–1519, 1997. 10

[32] Luc Blanchet. Gravitational Radiation from Post-Newtonian Sources and Inspiralling Compact Binaries. *Living Rev. Rel.*, 17:2, 2014. 16

[33] Nigel T. Bishop and Luciano Rezzolla. Extraction of Gravitational Waves in Numerical Relativity. *Living Rev. Rel.*, 19:2, 2016. 16

[34] Luis Lehner and Frans Pretorius. Numerical Relativity and Astrophysics. *Ann. Rev. Astron. Astrophys.*, 52:661–694, 2014. 16

[35] Ulrich Sperhake. The numerical relativity breakthrough for binary black holes. *Class. Quant. Grav.*, 32(12):124011, 2015. 16

[36] Chris L. Fryer and Kimberly C. B. New. Gravitational waves from gravitational collapse. *Living Rev. Rel.*, 14:1, 2011. 16

[37] N. Andersson and G. L. Comer. Probing neutron star superfluidity with gravitational wave data. *Phys. Rev. Lett.*, 87:241101, 2001. 16

[38] Anthony L. Piro and Eric Pfahl. Fragmentation of Collapsar Disks and the Production of Gravitational Waves. *Astrophys. J.*, 658:1173, 2007. 16

[39] Maurice H. P. M. van Putten. Gravitational waveforms of kerr black holes interacting with high-density matter. *The Astrophysical Journal*, 684(2):L91–L94, aug 2008. 16

[40] Alessandra Corsi and Benjamin J. Owen. Maximum gravitational-wave energy emissible in magnetar flares. *Phys. Rev. D*, 83:104014, 2011. 16

[41] Paul D. Lasky. Gravitational Waves from Neutron Stars: A Review. *Publ. Astron. Soc. Austral.*, 32:e034, 2015. 16

[42] Thibault Damour and Alexander Vilenkin. Gravitational wave bursts from cusps and kinks on cosmic strings. *Phys. Rev. D*, 64:064008, 2001. 16

[43] Thibault Damour and Alexander Vilenkin. Gravitational radiation from cosmic (super)strings: Bursts, stochastic background, and observational windows. *Phys. Rev. D*, 71:063510, 2005. 16, 18

[44] Magdalena Sieniawska and Michał Bejger. Continuous gravitational waves from neutron stars: current status and prospects. *Universe*, 5(11):217, 2019. 17

[45] Michael S. Turner. Detectability of inflation produced gravitational waves. *Phys. Rev. D*, 55:R435–R439, 1997. 17

[46] Richard Easther and Eugene A. Lim. Stochastic gravitational wave production after inflation. *JCAP*, 04:010, 2006. 17

[47] Richard Easther, John T. Giblin, Jr., and Eugene A. Lim. Gravitational Wave Production At The End Of Inflation. *Phys. Rev. Lett.*, 99:221301, 2007. 17

[48] Xing-Jiang Zhu, Eric J. Howell, David G. Blair, and Zong-Hong Zhu. On the gravitational wave background from compact binary coalescences in the band of ground-based interferometers. *Mon. Not. Roy. Astron. Soc.*, 431(1):882–899, 2013. 18

[49] C. Wu, V. Mandic, and T. Regimbau. Accessibility of the Gravitational-Wave Background due to Binary Coalescences to Second and Third Generation Gravitational-Wave Detectors. *Phys. Rev. D*, 85:104024, 2012. 18

[50] S. Marassi, R. Schneider, G. Corvino, V. Ferrari, and S. Portegies Zwart. Imprint of the merger and ring-down on the gravitational wave background from black hole binaries coalescence. *Phys. Rev. D*, 84:124037, 2011. 18

[51] Pablo A. Rosado. Gravitational wave background from binary systems. *Phys. Rev. D*, 84:084004, 2011. 18

[52] Xing-Jiang Zhu, Eric Howell, and David Blair. Observational upper limits on the gravitational wave production of core collapse supernovae. *Mon. Not. Roy. Astron. Soc.*, 409:L132–L136, 2010. 18

[53] Alessandra Buonanno, Gunter Sigl, Georg G. Raffelt, Hans-Thomas Janka, and Ewald Muller. Stochastic gravitational wave background from cosmological supernovae. *Phys. Rev. D*, 72:084001, 2005. 18

[54] T. Regimbau and Jose A. de Freitas Pacheco. Cosmic background of gravitational waves from rotating neutron stars. *Astron. Astrophys.*, 376:381, 2001. 18

[55] Xavier Siemens, Vuk Mandic, and Jolien Creighton. Gravitational wave stochastic background from cosmic (super)strings. *Phys. Rev. Lett.*, 98:111101, 2007. 18

[56] Luca Marzola, Antonio Racioppi, and Ville Vaskonen. Phase transition and gravitational wave phenomenology of scalar conformal extensions of the Standard Model. *Eur. Phys. J. C*, 77(7):484, 2017. 18

[57] Benedict Von Harling, Alex Pomarol, Oriol Pujolàs, and Fabrizio Rompineve. Peccei-Quinn Phase Transition at LIGO. *JHEP*, 04:195, 2020. 18

[58] Alberto Sesana, Francesco Haardt, Piero Madau, and Marta Volonteri. Low - frequency gravitational radiation from coalescing massive black hole binaries in hierarchical cosmologies. *Astrophys. J.*, 611:623–632, 2004. 18

[59] Seth E. Timpano, Louis J. Rubbo, and Neil J. Cornish. Characterizing the galactic gravitational wave background with LISA. *Phys. Rev. D*, 73:122001, 2006. 18

[60] T. Akutsu et al. Overview of KAGRA: Detector design and construction history. 5 2020. 19

[61] Katherine L. Dooley. Status of GEO 600. *J. Phys. Conf. Ser.*, 610(1):012015, 2015. 19

[62] Bala Iyer, Tarun Souradeep, C.S. Unnikrishnan, Sanjeev Dhurandhar, Sendhil Raja, Ajai Kumar, and Anand Sengupta. Ligo-india, proposal of the consortium for indian initiative in gravitational-wave observations (indigo). 19

[63] M. Saleem et al. The Science Case for LIGO-India. 5 2021. 19

[64] R. W. P. Drever, J. L. Hall, F. V. Kowalski, J. Hough, G. M. Ford, A. J. Munley, and H. Ward. Laser phase and frequency stabilization using an optical resonator. *Applied Physics B: Lasers and Optics*, 31(2):97–105, June 1983. 21

[65] B. P. Abbott et al. GW150914: The Advanced LIGO Detectors in the Era of First Discoveries. *Phys. Rev. Lett.*, 116(13):131103, 2016. 21

[66] F. Matichard et al. Seismic isolation of Advanced LIGO: Review of strategy, instrumentation and performance. *Class. Quant. Grav.*, 32(18):185003, 2015. 23

[67] V. B. Braginsky and F. Ya. Khalili. Quantum nondemolition measurements: the route from toys to tools. *Rev. Mod. Phys.*, 68:1–11, 1996. 24

[68] Carlton Caves. Quantum-Mechanical Radiation-Pressure Fluctuations in an Interferometer. *Phys. Rev. Lett.*, 45(2):75–79, 1980. 24

[69] C. M. Caves. Quantum Mechanical Noise in an Interferometer. *Phys. Rev. D*, 23:1693–1708, 1981. 24

[70] Haocun Yu et al. Quantum correlations between light and the kilogram-mass mirrors of LIGO. *Nature*, 583(7814):43–47, 2020. 24

[71] F. Acernese et al. Quantum Backaction on Kg-Scale Mirrors: Observation of Radiation Pressure Noise in the Advanced Virgo Detector. *Phys. Rev. Lett.*, 125(13):131101, 2020. 24

[72] F. Acernese et al. Increasing the Astrophysical Reach of the Advanced Virgo Detector via the Application of Squeezed Vacuum States of Light. *Phys. Rev. Lett.*, 123(23):231108, 2019. 24

[73] S. Karki et al. The Advanced LIGO Photon Calibrators. *Rev. Sci. Instrum.*, 87(11):114503, 2016. 26

[74] Ling Sun et al. Characterization of systematic error in Advanced LIGO calibration. *Class. Quant. Grav.*, 37(22):225008, 2020. 26

[75] Neil J. Cornish, Tyson B. Littenberg, Bence Bécsy, Katerina Chatziioannou, James A. Clark, Sudarshan Ghonge, and Margaret Millhouse. BayesWave analysis pipeline in the era of gravitational wave observations. *Phys. Rev. D*, 103(4):044006, 2021. 26

[76] John Zweizig and Keith Riles. Information on self-gating of h(t) used in O3 continuous-wave searches. https://dcc.ligo.org/T2000384/public. 26

[77] Andrew Matas, Irina Dvorkin, Alba Romero, and Tania Regimbau. Applying gating to stochastic searches in O3. https://dcc.ligo.org/P2000546/public. 26

[78] Patrick M. Meyers, Katarina Martinovic, Nelson Christensen, and Mairi Sakellariadou. Detecting a stochastic gravitational-wave background in the presence of correlated magnetic noise. *Phys. Rev. D*, 102(10):102005, 2020. 28

[79] R. Abbott et al. Upper Limits on the Isotropic Gravitational-Wave Background from Advanced LIGOs and Advanced Virgo's Third Observing Run. 1 2021. 28, 31, 50

[80] R. Abbott et al. GWTC-2: Compact Binary Coalescences Observed by LIGO and Virgo During the First Half of the Third Observing Run. 10 2020. 29, 37

[81] B. P. Abbott et al. GWTC-1: A Gravitational-Wave Transient Catalog of Compact Binary Mergers Observed by LIGO and Virgo during the First and Second Observing Runs. *Phys. Rev.*, X9(3):031040, 2019. 29, 37, 73, 87

[82] B. P. Abbott et al. GW190425: Observation of a Compact Binary Coalescence with Total Mass $\sim 3.4 M_{\odot}$. *Astrophys. J. Lett.*, 892(1):L3, 2020. 29

[83] R. Abbott et al. GW190521: A Binary Black Hole Merger with a Total Mass of $150 M_{\odot}$. *Phys. Rev. Lett.*, 125(10):101102, 2020. 29

[84] M. J. Graham et al. Candidate Electromagnetic Counterpart to the Binary Black Hole Merger Gravitational Wave Event S190521g. *Phys. Rev. Lett.*, 124(25):251102, 2020. 29, 46, 87

[85] Gregory Ashton, Kendall Ackley, Ignacio Magaña Hernandez, and Brandon Piotrzkowski. Current observations are insufficient to confidently associate the binary black hole merger GW190521 with AGN J124942.3+344929. 10 2020. 29

[86] Antonella Palmese, Maya Fishbach, Colin J. Burke, James T. Annis, and Xin Liu. Do LIGO/Virgo black hole mergers produce AGN flares? The case of GW190521 and prospects for reaching a confident association. 3 2021. 29

[87] R. Abbott et al. Population Properties of Compact Objects from the Second LIGO-Virgo Gravitational-Wave Transient Catalog. 10 2020. 29, 43

[88] B. P. Abbott et al. Gravitational Waves and Gamma-rays from a Binary Neutron Star Merger: GW170817 and GRB 170817A. *Astrophys. J. Lett.*, 848(2):L13, 2017. 30

[89] B. P. Abbott et al. GW170817: Measurements of neutron star radii and equation of state. *Phys. Rev. Lett.*, 121(16):161101, 2018. 30, 132

[90] Daniel J. Mortlock, Stephen M. Feeney, Hiranya V. Peiris, Andrew R. Williamson, and Samaya M. Nissanke. Unbiased Hubble constant estimation from binary neutron star mergers. *Phys. Rev. D*, 100(10):103523, 2019. 30

[91] Stephen M. Feeney, Hiranya V. Peiris, Andrew R. Williamson, Samaya M. Nissanke, Daniel J. Mortlock, Justin Alsing, and Dan Scolnic. Prospects for resolving the Hubble constant tension with standard sirens. *Phys. Rev. Lett.*, 122(6):061105, 2019. 30

[92] Hsin-Yu Chen, Maya Fishbach, and Daniel E. Holz. A two per cent Hubble constant measurement from standard sirens within five years. *Nature*, 562(7728):545–547, 2018. 30

[93] Joseph D. Romano and Neil J. Cornish. Detection methods for stochastic

gravitational-wave backgrounds: a unified treatment. *Living Rev. Rel.*, 20(1):2, 2017. 34, 35, 77, 87, 107

[94] Renate Meyer, Matthew C. Edwards, Patricio Maturana-Russel, and Nelson Christensen. Computational Techniques for Parameter Estimation of Gravitational Wave Signals. 9 2020. 35

[95] J. Veitch et al. Parameter estimation for compact binaries with ground-based gravitational-wave observations using the LALInference software library. *Phys. Rev. D*, 91(4):042003, 2015. 37, 42, 43, 77, 89

[96] Gregory Ashton et al. BILBY: A user-friendly Bayesian inference library for gravitational-wave astronomy. *Astrophys. J. Suppl.*, 241(2):27, 2019. 37, 89, 98

[97] I. M. Romero-Shaw et al. Bayesian inference for compact binary coalescences with bilby: validation and application to the first LIGO–Virgo gravitational-wave transient catalogue. *Mon. Not. Roy. Astron. Soc.*, 499(3):3295–3319, 2020. 37

[98] Neil J. Cornish and Joseph D. Romano. Towards a unified treatment of gravitational-wave data analysis. *Phys. Rev. D*, 87(12):122003, 2013. 37, 119

[99] Neil J. Cornish and Tyson B. Littenberg. BayesWave: Bayesian Inference for Gravitational Wave Bursts and Instrument Glitches. *Class. Quant. Grav.*, 32(13):135012, 2015. 39, 49, 74, 78, 129

[100] F. James. *Statistical Methods in Experimental Physics*. World Scientific, 2006. 39

[101] Thomas D. Abbott et al. Improved analysis of GW150914 using a fully spin-precessing waveform Model. *Phys. Rev. X*, 6(4):041014, 2016. 40

[102] Nicholas Metropolis, Arianna W. Rosenbluth, Marshall N. Rosenbluth, Augusta H. Teller, and Edward Teller. Equation of State Calculations by Fast Computing Machines. *jcp*, 21(6):1087–1092, June 1953. 41

[103] W. K. Hastings. Monte Carlo Sampling Methods Using Markov Chains and Their Applications. *Biometrika*, 57:97–109, 1970. 41

[104] S. Geman and D. Geman. Stochastic relaxation, gibbs distributions, and the bayesian restoration of images. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, PAMI-6(6):721–741, 1984. 42

[105] Simon Duane, A. D. Kennedy, Brian J. Pendleton, and Duncan Roweth. Hybrid Monte Carlo. *Physics Letters B*, 195(2):216–222, September 1987. 42

[106] Daniel Foreman-Mackey, David W. Hogg, Dustin Lang, and Jonathan Goodman. emcee: The MCMC hammer. *Publications of the Astronomical Society of the Pacific*, 125(925):306–312, mar 2013. 42

[107] John Salvatier, Thomas V. Wiecki, and Christopher Fonnesbeck. Probabilistic programming in Python using PyMC3. *PeerJ Computer Science*, 2:e55, April 2016. 42

[108] Stan Development Team. Stan User's Guide. 42

[109] John Skilling. Nested sampling for general Bayesian computation. *Bayesian Analysis*, 1(4):833–859, 2006. 42

[110] Joshua S Speagle. dynesty: a dynamic nested sampling package for estimating Bayesian posteriors and evidences. *Monthly Notices of the Royal Astronomical Society*, 493(3):3132–3158, 02 2020. 43, 98, 118

[111] F. Feroz, M. P. Hobson, and M. Bridges. MultiNest: an efficient and robust Bayesian inference tool for cosmology and particle physics. *Mon. Not. Roy. Astron. Soc.*, 398:1601–1614, 2009. 43, 80

[112] J. Buchner, A. Georgakakis, K. Nandra, L. Hsu, C. Rangel, M. Brightman, A. Merloni, M. Salvato, J. Donley, and D. Kocevski. X-ray spectral modelling of the AGN obscuring region in the CDFS: Bayesian model selection and catalogue. *Astronomy and Astrophysics*, 564:A125, April 2014. 43, 80

[113] Johannes Buchner. A statistical test for Nested Sampling algorithms. *arXiv e-prints*, page arXiv:1407.5459, July 2014. 43

[114] Johannes Buchner. Collaborative Nested Sampling: Big Data vs. complex physical models. *arXiv e-prints*, page arXiv:1707.04476, July 2017. 43

[115] B. P. Abbott et al. Binary Black Hole Population Properties Inferred from the First and Second Observing Runs of Advanced LIGO and Advanced Virgo. *Astrophys. J.*, 882(2):L24, 2019. 43, 87

[116] Salvatore Vitale, Davide Gerosa, Will M. Farr, and Stephen R. Taylor. Inferring the properties of a population of compact binaries in presence of selection effects. 7 2020. 44

[117] D. A. Coulter et al. Swope Supernova Survey 2017a (SSS17a), the Optical Counterpart to a Gravitational Wave Source. *Science*, 358:1556, 2017. 47

[118] N. R. Tanvir et al. The Emergence of a Lanthanide-Rich Kilonova Following the Merger of Two Neutron Stars. *Astrophys. J. Lett.*, 848(2):L27, 2017. 47

[119] Iair Arcavi et al. Optical emission from a kilonova following a gravitational-wave-detected neutron-star merger. *Nature*, 551:64, 2017. 47

[120] Thomas W. Baumgarte, Stuart L. Shapiro, and Masaru Shibata. On the maximum mass of differentially rotating neutron stars. *Astrophys. J. Lett.*, 528:L29, 2000. 47

[121] Luca Baiotti, Bruno Giacomazzo, and Luciano Rezzolla. Accurate evolutions of inspiralling neutron-star binaries: prompt and delayed collapse to black hole. *Phys. Rev. D*, 78:084033, 2008. 47

[122] Luca Baiotti and Luciano Rezzolla. Binary neutron star mergers: a review of Einstein's richest laboratory. *Rept. Prog. Phys.*, 80(9):096901, 2017. 47, 73

[123] A. Bauswein, T. W. Baumgarte, and H. T. Janka. Prompt merger collapse and the maximum mass of neutron stars. *Phys. Rev. Lett.*, 111(13):131101, 2013. 47

[124] Mark G. Alford, Luke Bovard, Matthias Hanauske, Luciano Rezzolla, and Kai Schwenzer. Viscous Dissipation and Heat Conduction in Binary Neutron-Star Mergers. *Phys. Rev. Lett.*, 120(4):041101, 2018. 48

[125] Riccardo Ciolfi, Wolfgang Kastaun, Jay Vijay Kalinani, and Bruno Giacomazzo. First 100 ms of a long-lived magnetized neutron star formed in a binary neutron star merger. *Phys. Rev. D*, 100(2):023005, 2019. 48

[126] Roberto De Pietri, Alessandra Feo, José A. Font, Frank Löffler, Francesco Maione, Michele Pasquali, and Nikolaos Stergioulas. Convective Excitation of Inertial Modes in Binary Neutron Star Mergers. *Phys. Rev. Lett.*, 120(22):221101, 2018. 48

[127] Elias R. Most, L. Jens Papenfort, Veronica Dexheimer, Matthias Hanauske, Stefan Schramm, Horst Stöcker, and Luciano Rezzolla. Signatures of quark-hadron phase transitions in general-relativistic neutron-star mergers. *Phys. Rev. Lett.*, 122(6):061101, 2019. 48

[128] Masaru Shibata, Kenta Kiuchi, and Yu-ichiro Sekiguchi. General relativistic viscous hydrodynamics of differentially rotating neutron stars. *Phys. Rev. D*, 95(8):083005, 2017. 48

[129] Daniel M. Siegel, Riccardo Ciolfi, Abraham I. Harte, and Luciano Rezzolla. Magnetorotational instability in relativistic hypermassive neutron stars. *Phys. Rev. D*, 87(12):121302, 2013. 48

[130] S. Klimenko et al. Method for detection and reconstruction of gravitational wave transients with networks of advanced detectors. *Phys. Rev. D*, 93(4):042004, 2016. 49

[131] B. P. Abbott et al. Search for Post-merger Gravitational Waves from the Remnant of the Binary Neutron Star Merger GW170817. *Astrophys. J. Lett.*, 851(1):L16, 2017. 49, 56, 69, 73, 74

[132] Eric Thrane et al. Long gravitational-wave transients and associated detection strategies for a network of terrestrial interferometers. *Phys. Rev. D*, 83:083004, 2011. 49, 50, 53, 55, 76

[133] S. Suvorova, L. Sun, A. Melatos, W. Moran, and R. J. Evans. Hidden Markov model tracking of continuous gravitational waves from a neutron star with wandering spin. *Phys. Rev. D*, 93(12):123009, 2016. 49

[134] L. Sun, A. Melatos, S. Suvorova, W. Moran, and R. J. Evans. Hidden Markov model tracking of continuous gravitational waves from young supernova remnants. *Phys. Rev. D*, 97(4):043013, 2018. 49

[135] Ling Sun and Andrew Melatos. Application of hidden Markov model tracking to the search for long-duration transient gravitational waves from the remnant of the binary neutron star merger GW170817. *Phys. Rev. D*, 99(12):123003, 2019. 49

[136] Badri Krishnan, Alicia M. Sintes, Maria Alessandra Papa, Bernard F. Schutz, Sergio Frasca, and Cristiano Palomba. The Hough transform search for continuous gravitational waves. *Phys. Rev. D*, 70:082001, 2004. 49

[137] Andrew Miller et al. Method to search for long duration gravitational wave transients from isolated neutron stars using the generalized frequency-Hough transform. *Phys. Rev. D*, 98(10):102004, 2018. 49

[138] Miquel Oliver, David Keitel, and Alicia M. Sintes. Adaptive transient Hough method for long-duration gravitational wave transients. *Phys. Rev. D*, 99(10):104067, 2019. 49

[139] B. P. Abbott et al. Upper Limits on Gravitational Waves from Scorpius X-1 from a Model-Based Cross-Correlation Search in Advanced LIGO Data. *Astrophys. J.*, 847(1):47, 2017. 50

[140] R. Abbott et al. Search for anisotropic gravitational-wave backgrounds using data from Advanced LIGO's and Advanced Virgo's first three observing runs. 3 2021. 50

[141] Shivaraj Kandhasamy. Searches for stochastic gravitational waves and long gravitational wave transients in ligo s5 data, 2013. 53

[142] Rubab Khan and Shourov Chatterji. Enhancing the capabilities of LIGO time-frequency plane searches through clustering. *Class. Quant. Grav.*, 26:155009, 2009. 55, 76

[143] Tanner Prestegard. Unmodeled searches for long-lasting gravitational-wave signals with ligo and studies of underground seismic noise for future gravitational-wave detectors, 2016. 55

[144] M. Coughlin, E. Thrane, and N. Christensen. Detecting compact binary coalescences with seedless clustering. *Phys. Rev. D*, 90(8):083005, 2014. 55, 76

[145] Eric Thrane and Michael Coughlin. Searching for gravitational-wave transients with a qualitative signal model: seedless clustering strategies. *Phys. Rev. D*, 88(8):083010, 2013. 55, 56, 76

[146] Eric Thrane and Michael Coughlin. Seedless clustering in all-sky searches for gravitational-wave transients. *Phys. Rev. D*, 89(6):063012, 2014. 55, 56, 76

[147] B. P. Abbott et al. All-sky search for long-duration gravitational wave transients with initial LIGO. *Phys. Rev. D*, 93(4):042005, 2016. 56

[148] B. P. Abbott et al. All-sky search for long-duration gravitational-wave transients in the second Advanced LIGO observing run. *Phys. Rev. D*, 99(10):104033, 2019. 56

[149] Benjamin P. Abbott et al. All-sky search for long-duration gravitational wave transients in the first Advanced LIGO observing run. *Class. Quant. Grav.*, 35(6):065009, 2018. 56

[150] B. P. Abbott et al. Search for Transient Gravitational-wave Signals Associated with Magnetar Bursts during Advanced LIGOs Second Observing Run. *Astrophys. J.*, 874(2):163, 2019. 56

[151] Sharan Banagiri, Ling Sun, Michael W. Coughlin, and Andrew Melatos. Search strategies for long gravitational-wave transients: hidden Markov model tracking and seedless clustering. *Phys. Rev. D*, 100(2):024034, 2019. 56, 71

[152] Eric Thrane, Vuk Mandic, and Nelson Christensen. Detecting very long-lived gravitational-wave transients lasting hours to weeks. *Phys. Rev. D*, 91(10):104021, 2015. 58

[153] D. Davis, T. J. Massinger, A. P. Lundgren, J. C. Driggers, A. L. Urban, and L. K. Nuttall. Improving the Sensitivity of Advanced LIGO Using Noise Subtraction. *Class. Quant. Grav.*, 36(5):055011, 2019. 59

[154] Tanner Prestegard, Eric Thrane, Nelson L. Christensen, Michael W. Coughlin,

Ben Hubbert, Shivaraj Kandhasamy, Evan MacAyeal, and Vuk Mandic. Identification of noise artifacts in searches for long-duration gravitational-wave transients. *Class. Quant. Grav.*, 29:095018, 2012. 60

[155] Nikhil Sarin, Paul D. Lasky, Letizia Sammut, and Greg Ashton. X-ray guided gravitational-wave search for binary neutron star merger remnants. *Phys. Rev. D*, 98(4):043011, 2018. 65, 66, 73, 74

[156] S.L. Shapiro and S.A. Teukolsky. *Black Holes, White Dwarfs, and Neutron Stars: The Physics of Compact Objects*. Wiley, 2008. 65

[157] S. Bonazzola and E. Gourgoulhon. Gravitational waves from pulsars: Emission by the magnetic field induced distortion. *Astron. Astrophys.*, 312:675, 1996. 65

[158] M. Punturo et al. The Einstein Telescope: A third-generation gravitational wave observatory. *Class. Quant. Grav.*, 27:194002, 2010. 70

[159] Benjamin P Abbott et al. Exploring the Sensitivity of Next Generation Gravitational Wave Detectors. *Class. Quant. Grav.*, 34(4):044001, 2017. 70

[160] K. Ackley et al. Neutron Star Extreme Matter Observatory: A kilohertz-band gravitational-wave detector in the global network. *Publ. Astron. Soc. Austral.*, 37:e047, 2020. 70

[161] B. P. Abbott et al. GW170814: A Three-Detector Observation of Gravitational Waves from a Binary Black Hole Coalescence. *Phys. Rev. Lett.*, 119(14):141101, 2017. 73

[162] Benjamin P Abbott et al. Model comparison from LIGO–Virgo data on GW170817's binary components and consequences for the merger remnant. *Class. Quant. Grav.*, 37(4):045006, 2020. 73

[163] Maurice H. P. M. van Putten and Massimo Della Valle. Observational evidence for Extended Emission to GW170817. *Mon. Not. Roy. Astron. Soc.*, 482(1):L46–L49, 2019. 73

[164] Anthony L. Piro, Bruno Giacomazzo, and Rosalba Perna. The Fate of Neutron Star Binary Mergers. *Astrophys. J. Lett.*, 844(2):L19, 2017. 73

[165] A. Rowlinson, P. T. O'Brien, B. D. Metzger, N. R. Tanvir, and A. J. Levan. Signatures of magnetar central engines in short GRB lightcurves. *Mon. Not. Roy. Astron. Soc.*, 430:1061, 2013. 73

[166] Simone Dall'Osso, Bruno Giacomazzo, Rosalba Perna, and Luigi Stella. Gravitational waves from massive magnetars formed in binary neutron star mergers. *Astrophys. J.*, 798(1):25, 2015. 73, 74

[167] M. Punturo et al. The third generation of gravitational wave observatories and their science reach. *Class. Quant. Grav.*, 27:084007, 2010. 73

[168] S. Hild et al. Sensitivity Studies for Third-Generation Gravitational Wave Observatories. *Class. Quant. Grav.*, 28:094013, 2011. 73

[169] Curt Cutler. Gravitational waves from neutron stars with large toroidal B fields. *Phys. Rev. D*, 66:084025, 2002. 74

[170] Daniela D. Doneva, Kostas D. Kokkotas, and Pantelis Pnigouras. Gravitational wave afterglow in binary neutron star mergers. *Phys. Rev. D*, 92(10):104040, 2015. 74

[171] Paul D Lasky and Kostas Glampedakis. Observationally constraining gravitational wave emission from short gamma-ray burst remnants. *Mon. Not. Roy. Astron. Soc.*, 458(2):1660–1670, 2016. 74

[172] Katerina Chatziioannou, James Alexander Clark, Andreas Bauswein, Margaret Millhouse, Tyson B. Littenberg, and Neil Cornish. Inferring the post-merger gravitational wave emission from binary neutron star coalescences. *Phys. Rev. D*, 96(12):124035, 2017. 74

[173] Paul D. Lasky, Cristiano Leris, Antonia Rowlinson, and Kostas Glampedakis. The braking index of millisecond magnetars. *Astrophys. J. Lett.*, 843(1):L1, 2017. 74, 84

[174] Warren G. Anderson, Patrick R. Brady, Jolien D. E. Creighton, and Eanna E. Flanagan. An Excess power statistic for detection of burst sources of gravitational radiation. *Phys. Rev. D*, 63:042003, 2001. 76

[175] Patrick J. Sutton et al. X-Pipeline: An Analysis package for autonomous gravitational-wave burst searches. *New J. Phys.*, 12:053034, 2010. 76

[176] S. Klimenko, I. Yakushin, A. Mercer, and Guenakh Mitselmakher. Coherent method for detection of gravitational wave bursts. *Class. Quant. Grav.*, 25:114029, 2008. 76

[177] Prestegard and Thrane. LIGO-l1200204-v1: Burstegard: a hierarchical clustering algorithm. 76

[178] Tyson B. Littenberg and Neil J. Cornish. Separating Gravitational Wave Signals from Instrument Artifacts. *Phys. Rev. D*, 82:103007, 2010. 76

[179] George Arfken. *Mathematical Methods for Physicists*. Elsevier, 2012. 77, 114

[180] Christian Rover, Renate Meyer, and Nelson Christensen. Modelling coloured residual noise in gravitational-wave signal processing. *Class. Quant. Grav.*, 28:015010, 2011. 78

[181] Tyson B. Littenberg, Michael Coughlin, Benjamin Farr, and Will M. Farr. Fortifying the characterization of binary mergers in LIGO data. *Phys. Rev. D*, 88(8):084044, 2013. 78

[182] Kentaro Takami, Luciano Rezzolla, and Luca Baiotti. Constraining the Equation of State of Neutron Stars from Binary Mergers. *Phys. Rev. Lett.*, 113(9):091104, 2014. 84

[183] Sebastiano Bernuzzi, Tim Dietrich, and Alessandro Nagar. Modeling the complete gravitational wave spectrum of neutron star mergers. *Phys. Rev. Lett.*, 115(9):091101, 2015. 84

[184] Ka Wa Tsang, Tim Dietrich, and Chris Van Den Broeck. Modeling the post-merger gravitational wave signal and extracting binary properties from future binary neutron star detections. *Phys. Rev. D*, 100(4):044047, 2019. 84

[185] Rory Smith and Eric Thrane. Optimal Search for an Astrophysical Gravitational-Wave Background. *Phys. Rev.*, X8(2):021019, 2018. 87, 88, 95, 101, 104

[186] V. Connaughton et al. Fermi GBM Observations of LIGO Gravitational Wave event GW150914. *Astrophys. J. Lett.*, 826(1):L6, 2016. 87

[187] https://gracedb.ligo.org/superevents/public/O3/. 87

[188] Alexander H. Nitz, Thomas Dent, Gareth S. Davies, Sumit Kumar, Collin D. Capano, Ian Harry, Simone Mozzon, Laura Nuttall, Andrew Lundgren, and Márton Tápai. 2-OGC: Open Gravitational-wave Catalog of binary mergers from analysis of public Advanced LIGO and Virgo data. 2019. 87

[189] Tejaswi Venumadhav, Barak Zackay, Javier Roulet, Liang Dai, and Matias Zaldarriaga. New Binary Black Hole Mergers in the Second Observing Run of Advanced LIGO and Advanced Virgo. 2019. 87

[190] Tejaswi Venumadhav, Barak Zackay, Javier Roulet, Liang Dai, and Matias Zaldarriaga. New search pipeline for compact binary mergers: Results for binary black holes in the first observing run of Advanced LIGO. *Phys. Rev.*, D100(2):023011, 2019. 87

[191] Richard Stiskalek, John Veitch, and Chris Messenger. Are stellar mass binary black hole mergers isotropically distributed? 2020. 87, 95

[192] Ethan Payne, Sharan Banagiri, Paul Lasky, and Eric Thrane. Searching for anisotropy in the distribution of binary black hole mergers. 6 2020. 87, 104, 109, 116

[193] B. P. Abbott et al. Search for the isotropic stochastic background using data from Advanced LIGOs second observing run. *Phys. Rev.*, D100(6):061101, 2019. 87

[194] Bruce Allen and Joseph D. Romano. Detecting a stochastic background of gravitational radiation: Signal processing strategies and sensitivities. *Phys. Rev.*, D59:102001, 1999. 87, 110

[195] Eric Thrane, Stefan Ballmer, Joseph D. Romano, Sanjit Mitra, Dipongkar Talukder, Sukanta Bose, and Vuk Mandic. Probing the anisotropies of a stochastic gravitational-wave background using a network of ground-based laser interferometers. *Phys. Rev. D*, 80:122002, Dec 2009. 87, 108, 109

[196] Benjamin P. Abbott et al. Directional Limits on Persistent Gravitational Waves from Advanced LIGOs First Observing Run. *Phys. Rev. Lett.*, 118(12):121102, 2017. 87

[197] B. P. Abbott et al. Directional limits on persistent gravitational waves using data from Advanced LIGOs first two observing runs. *Phys. Rev.*, D100(6):062001, 2019. 87, 108

[198] Arianna Renzini and Carlo Contaldi. Improved limits on a stochastic gravitational-wave background and its anisotropies from Advanced LIGO O1 and O2 runs. *Phys. Rev.*, D100(6):063527, 2019. 87

[199] Anirban Ain, Jishnu Suresh, and Sanjit Mitra. Very fast stochastic gravitational wave background map making using folded data. *Phys. Rev.*, D98(2):024001, 2018. 87

[200] Giulia Cusin, Irina Dvorkin, Cyril Pitrou, and Jean-Philippe Uzan. First predictions of the angular power spectrum of the astrophysical gravitational wave background. *Phys. Rev. Lett.*, 120:231101, 2018. 87

[201] Alexander C. Jenkins, Richard O'Shaughnessy, Mairi Sakellariadou, and Daniel Wysocki. Anisotropies in the astrophysical gravitational-wave background: The impact of black hole distributions. *Phys. Rev. Lett.*, 122(11):111101, 2019. 87

[202] Alexander C. Jenkins, Mairi Sakellariadou, Tania Regimbau, and Eric Slezak. Anisotropies in the astrophysical gravitational-wave background: Predictions for the detection of compact binaries by LIGO and Virgo. *Phys. Rev.*, D98(6):063501, 2018. 87

[203] Daniele Bertacca, Angelo Ricciardone, Nicola Bellomo, Alexander C. Jenkins, Sabino Matarrese, Alvise Raccanelli, Tania Regimbau, and Mairi Sakellariadou. Projection effects on the observed angular spectrum of the astrophysical stochastic gravitational wave background. 2019. 87

[204] Giulia Cusin, Cyril Pitrou, and Jean-Philippe Uzan. Anisotropy of the astrophysical gravitational wave background: Analytic expression of the angular power spectrum and correlation with cosmological observations. *Phys. Rev.*, D96(10):103019, 2017. 87

[205] Giulia Cusin, Cyril Pitrou, and Jean-Philippe Uzan. The signal of the gravitational wave background and the angular correlation of its energy density. *Phys. Rev.*, D97(12):123527, 2018. 87

[206] Giulia Cusin, Irina Dvorkin, Cyril Pitrou, and Jean-Philippe Uzan. Properties of the stochastic astrophysical gravitational wave background: astrophysical sources dependencies. *Phys. Rev.*, D100(6):063004, 2019. 87

[207] Cyril Pitrou, Giulia Cusin, and Jean-Philippe Uzan. A unified view of anisotropies in the astrophysical gravitational wave background. 2019. 87

[208] Carlo R. Contaldi. Anisotropies of Gravitational Wave Backgrounds: A Line Of Sight Approach. *Phys. Lett.*, B771:9–12, 2017. 87

[209] Alexander C. Jenkins and Mairi Sakellariadou. Shot noise in the astrophysical gravitational-wave background. *Phys. Rev.*, D100(6):063508, 2019. 88

[210] Alexander C. Jenkins, Joseph D. Romano, and Mairi Sakellariadou. Estimating the angular power spectrum of the gravitational-wave background in the presence of shot noise. *Phys. Rev.*, D100(8):083501, 2019. 88

[211] Rory Smith, Colm Talbot, Francisco Hernandez Vivanco, and Eric Thrane. Inferring the population properties of binary black holes from unresolved gravitational waves. 4 2020. 88, 95, 104

[212] P.J.E. Peebles. *The Large-scale Structure of the Universe*. Princeton Series in Physics. Princeton University Press, 1980. 89, 90, 92

[213] Eric Thrane and Colm Talbot. An introduction to Bayesian inference in gravitational-wave astronomy: parameter estimation, model selection, and hierarchical models. *Publ. Astron. Soc. Austral.*, 36:e010, 2019. [Erratum: Publ.Astron.Soc.Austral. 37, e036 (2020)]. 89, 91

[214] Benjamin P Abbott et al. A guide to LIGO–Virgo detector noise and extraction of transient gravitational-wave signals. *Class. Quant. Grav.*, 37(5):055002, 2020. 89

[215] J.A. Peacock. *Cosmological Physics*. Cambridge Astrophysics. Cambridge University Press, 1999. 92

[216] M. R. Blanton, M. A. Bershady, B. Abolfathi, F. D. Albareti, C. Allende Prieto, A. Almeida, J. Alonso-García, F. Anders, S. F. Anderson, B. Andrews, and et al. Sloan digital sky survey iv: Mapping the milky way, nearby galaxies, and the distant universe. 2017. 93

[217] Hsin-Yu Chen, Reed Essick, Salvatore Vitale, Daniel E. Holz, and Erik Katsavounidis. Observational Selection Effects with Ground-based Gravitational Wave Detectors. *Astrophys. J.*, 835(1):31, 2017. 95

[218] B.P. Abbott et al. A gravitational-wave measurement of the Hubble constant following the second observing run of Advanced LIGO and Virgo. 8 2019. 95

[219] Patricia Schmidt, Mark Hannam, and Sascha Husa. Towards models of gravitational waveforms from generic binaries: A simple approximate mapping between precessing and non-precessing inspiral signals. *Phys. Rev. D*, 86:104063, 2012. 98

[220] Sebastian Khan, Sascha Husa, Mark Hannam, Frank Ohme, Michael Pürrer, Xisco Jiménez Forteza, and Alejandro Bohé. Frequency-domain gravitational waves from nonprecessing black-hole binaries. II. A phenomenological model for the advanced detector era. *Phys. Rev. D*, 93(4):044007, 2016. 98

[221] Andrea Zonca, Leo Singer, Daniel Lenz, Martin Reinecke, Cyrille Rosset, Eric Hivon, and Krzysztof Gorski. healpy: equal area pixelization and spherical harmonics transforms for data on the sphere in python. *Journal of Open Source Software*, 4(35):1298, March 2019. 98

[222] K. M. Górski, E. Hivon, A. J. Banday, B. D. Wandelt, F. K. Hansen, M. Reinecke, and M. Bartelmann. HEALPix: A Framework for High-Resolution Discretization and Fast Analysis of Data Distributed on the Sphere. *apj*, 622:759–771, April 2005. 98

[223] S. Dodelson and Academic Press (Londyn ; 1941-1969). *Modern Cosmology*. Elsevier Science, 2003. 103

[224] Suvodip Mukherjee and Benjamin D. Wandelt. Beyond the classical distance-redshift test: cross-correlating redshift-free standard candles and sirens with redshift surveys. 8 2018. 105

[225] Suvodip Mukherjee and Joseph Silk. Time-dependence of the astrophysical stochastic gravitational wave background. *Mon. Not. Roy. Astron. Soc.*, 491(4):4690–4701, 2020. 105

[226] Suvodip Mukherjee, Benjamin D. Wandelt, and Joseph Silk. Probing the theory of gravity with gravitational lensing of gravitational waves and galaxy surveys. *Mon. Not. Roy. Astron. Soc.*, 494(2):1956–1970, 2020. 105

[227] Kate Ziyan Yang, Vuk Mandic, Claudia Scarlata, and Sharan Banagiri. Searching for Cross-Correlation Between Stochastic Gravitational Wave Background and Galaxy Number Counts. 7 2020. 105

[228] Guadalupe Cañas-Herrera, Omar Contigiani, and Valeri Vardanyan. Cross-correlation of the astrophysical gravitational-wave background with galaxy clustering. 10 2019. 105

[229] David Alonso, Giulia Cusin, Pedro G. Ferreira, and Cyril Pitrou. Detecting the anisotropic astrophysical gravitational wave background in the presence of shot noise through cross-correlations. 2 2020. 105

[230] Aditya Vijaykumar, M.V.S. Saketh, Sumit Kumar, Parameswaran Ajith, and Tirthankar Roy Choudhury. Probing the large scale structure using gravitational-wave observations of binary black holes. 5 2020. 105

[231] Pau Amaro-Seoane. et al. Laser interferometer space antenna. *arXiv e-prints*, February 2017. 107, 116

[232] T R Marsh. Double white dwarfs and LISA. *Classical and Quantum Gravity*, 28(9):094019, apr 2011. 107

[233] Valeriya Korol, Elena M. Rossi, Paul J. Groot, Gijs Nelemans, Silvia Toonen, and Anthony G.A. Brown. Prospects for detection of detached double white dwarf binaries with Gaia, LSST and LISA. *Mon. Not. Roy. Astron. Soc.*, 470(2):1894–1910, 2017. 107

[234] Valeriya Korol, Orlin Koop, and Elena M. Rossi. Detectability of double white dwarfs in the Local Group with LISA. *Astrophys. J. Lett.*, 866(2):L20, 2018. 107

[235] V. Korol et al. Populations of double white dwarfs in Milky Way satellites and their detectability with LISA. *Astron. Astrophys.*, 638:A153, 2020. 107

[236] Pau Amaro-Seoane, Jonathan R. Gair, Marc Freitag, M. Coleman Miller, Ilya Mandel, Curt J. Cutler, and Stanislav Babak. Astrophysics, detection and science applications of intermediate- and extreme mass-ratio inspirals. *Class. Quant. Grav.*, 24:R113–R169, 2007. 107

[237] Jonathan R. Gair, Stanislav Babak, Alberto Sesana, Pau Amaro-Seoane, Enrico Barausse, Christopher P.L. Berry, Emanuele Berti, and Carlos Sopuerta. Prospects for observing extreme-mass-ratio inspirals with LISA. *J. Phys. Conf. Ser.*, 840(1):012021, 2017. 107

[238] Stanislav Babak, Jonathan Gair, Alberto Sesana, Enrico Barausse, Carlos F. Sopuerta, Christopher P.L. Berry, Emanuele Berti, Pau Amaro-Seoane, Antoine Petiteau, and Antoine Klein. Science with the space-based interferometer LISA. V: Extreme mass-ratio inspirals. *Phys. Rev. D*, 95(10):103012, 2017. 107

[239] Enrico Barausse, Jillian Bellovary, Emanuele Berti, Kelly Holley-Bockelmann, Brian Farris, Bangalore Sathyaprakash, and Alberto Sesana. Massive Black Hole Science with eLISA. *J. Phys. Conf. Ser.*, 610(1):012001, 2015. 107

[240] Antoine Klein et al. Science with the space-based interferometer eLISA: Supermassive black hole binaries. *Phys. Rev. D*, 93(2):024003, 2016. 107

[241] Mike Y.M. Lau, Ilya Mandel, Alejandro Vigna-Gómez, Coenraad J. Neijssel, Simon Stevenson, and Alberto Sesana. Detecting Double Neutron Stars with LISA. *Mon. Not. Roy. Astron. Soc.*, 492(3):3061–3072, 2020. 107

[242] Nicola Tamanini and Camilla Danielski. The gravitational-wave detection of exoplanets orbiting white dwarf binaries using lisa. *Nature Astronomy*, 3(9):858–866, Sep 2019. 107

[243] C. Danielski, V. Korol, N. Tamanini, and E.M. Rossi. Circumbinary exoplanets and brown dwarfs with the Laser Interferometer Space Antenna. *Astron. Astrophys.*, 632:A113, 2019. 107

[244] Pierre Auclair et al. Probing the gravitational wave background from cosmic strings with LISA. *JCAP*, 04:034, 2020. 107

[245] Chiara Caprini et al. Science with the space-based interferometer eLISA. II: Gravitational waves from cosmological phase transitions. *JCAP*, 04:001, 2016. 107

[246] Nicola Bartolo et al. Science with the space-based interferometer LISA. IV: Probing inflation with gravitational waves. *JCAP*, 12:026, 2016. 107

[247] Margot Fitz Axen, Sharan Banagiri, Andrew Matas, Chiara Caprini, and Vuk Mandic. Multiwavelength observations of cosmological phase transitions using LISA and Cosmic Explorer. *Phys. Rev. D*, 98(10):103508, 2018. 107

[248] Tania Regimbau. The astrophysical gravitational wave stochastic background. *Res. Astron. Astrophys.*, 11:369–390, 2011. 107, 109

[249] A. Brazier et al. The NANOGrav Program for Gravitational Waves and Fundamental Physics. 8 2019. 108

[250] R. N. Manchester et al. The Parkes Pulsar Timing Array Project. *Publ. Astron. Soc. Austral.*, 30:17, 2013. 108

[251] Michael Kramer and David J. Champion. The European Pulsar Timing Array and the Large European Array for Pulsars. *Class. Quant. Grav.*, 30:224009, 2013. 108

[252] B.P. Abbott et al. Search for the isotropic stochastic background using data from Advanced LIGOs second observing run. *Phys. Rev. D*, 100(6):061101, 2019. 108

[253] Zaven Arzoumanian et al. The NANOGrav 12.5 yr Data Set: Search for an Isotropic Stochastic Gravitational-wave Background. *Astrophys. J. Lett.*, 905(2):L34, 2020. 108

[254] Matthew Benacquista and K. Holley-Bockelmann. Consequences of disk scale height on LISA confusion noise from close white dwarf binaries. *Astrophys. J.*, 645:589–596, 2006. 108

[255] Katelyn Breivik, Chiara M. F. Mingarelli, and Shane L. Larson. Constraining Galactic Structure with the LISA White Dwarf Foreground. *Astrophys. J.*, 901(1):4, 2020. 108, 125, 128

[256] Carlo Ungarelli and Alberto Vecchio. Studying the anisotropy of the gravitational wave stochastic background with LISA. *Phys. Rev. D*, 64:121501, 2001. 108

[257] Hideaki Kudoh and Atsushi Taruya. Probing anisotropies of gravitational-wave backgrounds with a space-based interferometer: Geometric properties of antenna patterns and their angular power. *Phys. Rev. D*, 71:024025, 2005. 108

[258] Atsushi Taruya and Hideaki Kudoh. Probing anisotropies of gravitational-wave backgrounds with a space-based interferometer. II. Perturbative reconstruction of a low-frequency skymap. *Phys. Rev. D*, 72:104015, 2005. 108

[259] Atsushi Taruya. Probing anisotropies of gravitational-wave backgrounds with a space-based interferometer III: Reconstruction of a high-frequency skymap. *Phys. Rev. D*, 74:104022, 2006. 108

[260] A.I. Renzini and C.R. Contaldi. Mapping Incoherent Gravitational Wave Backgrounds. *Mon. Not. Roy. Astron. Soc.*, 481(4):4650–4661, 2018. 108

[261] Carlo R. Contaldi, Mauro Pieroni, Arianna I. Renzini, Giulia Cusin, Nikos Karnesis, Marco Peloso, Angelo Ricciardone, and Gianmassimo Tasinato. Maximum likelihood map-making with the Laser Interferometer Space Antenna. *Phys. Rev. D*, 102(4):043502, 2020. 108, 109

[262] Tyson Littenberg, Neil Cornish, Kristen Lackeos, and Travis Robson. Global

Analysis of the Gravitational Wave Signal from Galactic Binaries. *Phys. Rev. D*, 101(12):123021, 2020. 108

[263] Stephen R. Taylor, Rutger van Haasteren, and Alberto Sesana. From Bright Binaries To Bumpy Backgrounds: Mapping Realistic Gravitational Wave Skies With Pulsar-Timing Arrays. *Phys. Rev. D*, 102(8):084039, 2020. 109, 116

[264] E.S. Phinney. A Practical theorem on gravitational wave backgrounds. 7 2001. 111

[265] Neil J. Cornish. Detecting a stochastic gravitational wave background with the Laser Interferometer Space Antenna. *Phys. Rev. D*, 65:022004, 2002. 112

[266] Roland Schilling. Angular and frequency response of LISA. *Classical and Quantum Gravity*, 14(6):1513–1519, jun 1997. 112

[267] Stephen R. Taylor and Jonathan R. Gair. Searching For Anisotropic Gravitational-wave Backgrounds Using Pulsar Timing Arrays. *Phys. Rev. D*, 88:084001, 2013. 113

[268] S. R. Taylor et al. Limits on anisotropy in the nanohertz stochastic gravitational-wave background. *Phys. Rev. Lett.*, 115(4):041101, 2015. 113

[269] N. J. Cornish. Making maps with LISA. *Class. Quant. Grav.*, 19:1279–1283, 2002. 114

[270] Neil J. Cornish. Mapping the gravitational wave background. *Class. Quant. Grav.*, 18:4277–4292, 2001. 114

[271] Massimo Tinto and Sanjeev V. Dhurandhar. Time-delay interferometry. *Living Rev. Rel.*, 8:4, 2005. 117

[272] Matthew R. Adams and Neil J. Cornish. Discriminating between a Stochastic Gravitational Wave Background and Instrument Noise. *Phys. Rev. D*, 82:022002, 2010. 117, 119

[273] Daniel Foreman-Mackey, David W. Hogg, Dustin Lang, and Jonathan Goodman. emcee: The MCMC hammer. *Publications of the Astronomical Society of the Pacific*, 125(925):306–312, mar 2013. 118

[274] Aaron Meurer, Christopher P. Smith, Mateusz Paprocki, Ondřej Čertík, Sergey B. Kirpichev, Matthew Rocklin, AMiT Kumar, Sergiu Ivanov, Jason K. Moore, Sartaj Singh, Thilina Rathnayake, Sean Vig, Brian E. Granger, Richard P. Muller, Francesco Bonazzi, Harsh Gupta, Shivam Vats, Fredrik Johansson, Fabian Pedregosa, Matthew J. Curry, Andy R. Terrel, Štěpán Roučka, Ashutosh Saboo, Isuru Fernando, Sumith Kulal, Robert Cimrman, and Anthony Scopatz. Sympy: symbolic computing in python. *PeerJ Computer Science*, 3:e103, January 2017. 118

[275] Katelyn Breivik et al. COSMIC Variance in Binary Population Synthesis. *Astrophys. J.*, 898(1):71, 2020. 125

[276] Paul J. McMillan. Mass models of the Milky Way. *Monthly Notices of the Royal Astronomical Society*, 414(3):2446–2457, 06 2011. 125

[277] Matthew R. Adams and Neil J. Cornish. Detecting a Stochastic Gravitational Wave Background in the presence of a Galactic Foreground and Instrument Noise. *Phys. Rev. D*, 89(2):022001, 2014. 129

[278] Piero Madau and Tassos Fragos. Radiation Backgrounds at Cosmic Dawn: X-Rays from Compact Binaries. *Astrophys. J.*, 840(1):39, 2017. 131

[279] Thomas Callister, Maya Fishbach, Daniel Holz, and Will Farr. Shouts and Murmurs: Combining Individual Gravitational-Wave Sources with the Stochastic Background to Measure the History of Binary Black Hole Mergers. *Astrophys. J. Lett.*, 896(2):L32, 2020. 131

[280] S. Libanore, M. C. Artale, D. Karagiannis, M. Liguori, N. Bartolo, Y. Bouffanais, N. Giacobbo, M. Mapelli, and S. Matarrese. Gravitational Wave mergers as tracers of Large Scale Structures. *JCAP*, 02:035, 2021. 131

[281] Juan Calderón Bustillo, Tim Dietrich, Paul D. Lasky, and Samson H. W. Leong. Mapping the Universe Expansion: Enabling percent-level measurements of the Hubble Constant with a single binary neutron-star merger detection. 6 2020. 132

[282] Alvise Raccanelli, Ely D. Kovetz, Simeon Bird, Ilias Cholis, and Julian B. Munoz. Determining the progenitors of merging black-hole binaries. *Phys. Rev. D*, 94(2):023516, 2016. 132

[283] Giulio Scelfo, Nicola Bellomo, Alvise Raccanelli, Sabino Matarrese, and Licia Verde. GW×LSS: chasing the progenitors of merging binary black holes. *JCAP*, 09:039, 2018. 132

[284] Bendat and Piersol. *Random data : analysis and measurement procedures*. Wiley series in probability and statistics. Texts and references section. Wiley, New York, 3rd ed.. edition, 2000. 167

Appendix A

STAMP limits for the GW170817 long-lived post-merger search

Table A.1: The table lists the spin-down model injections that were done with the STAMP pipeline. From left to right, the parameters are the inclination, the starting frequency, the spin-down timescale, the eccentricity, the 90% distance sensitivity limit and the 90% energy limit. The data in this table is reproduced from Ref. [3] with permission.

$\cos\iota$	f_0 [Hz]	τ [s]	ϵ	$d^{90\%}$ [Mpc]	$E^{90\%}$ [$M_\odot c^2$]
0.0	500.0	100.0	0.0733	0.142	239.0
0.0	750.0	100.0	0.0326	0.055	3520.0
0.0	1000.0	100.0	0.0183	0.124	1250.0
0.0	1250.0	100.0	0.0117	0.12	2100.0
0.0	1500.0	100.0	0.00814	0.09	5360.0
0.0	1750.0	100.0	0.00598	0.057	18400.0
0.0	2000.0	100.0	0.00458	0.055	25100.0
0.0	500.0	1000.0	0.0232	0.082	711.0
0.0	750.0	1000.0	0.0103	0.068	2340.0

Continued on next page

Table A.1 – continued from previous page

$\cos\iota$	f_0 [Hz]	τ [s]	ϵ	$d^{90\%}$ [Mpc]	$E^{90\%}$ [$M_\odot c^2$]
0.0	1000.0	1000.0	0.00579	0.05	7730.0
0.0	1250.0	1000.0	0.00371	0.077	4990.0
0.0	1500.0	1000.0	0.00257	0.059	12500.0
0.0	1750.0	1000.0	0.00189	0.038	41400.0
0.0	2000.0	1000.0	0.00145	0.027	104000.0
0.0	2500.0	1000.0	0.000927	0.002	23400000.0
0.0	2750.0	1000.0	0.000766	0.004	11100000.0
0.0	3000.0	1000.0	0.000644	0.006	5320000.0
0.0	500.0	10000.0	0.00733	0.088	612.0
0.0	750.0	10000.0	0.00326	0.072	2060.0
0.0	1000.0	10000.0	0.00183	0.037	14100.0
0.0	1250.0	10000.0	0.00117	0.03	32900.0
0.0	1500.0	10000.0	0.000814	0.027	57500.0
0.0	1750.0	10000.0	0.000598	0.02	141000.0
0.0	2000.0	10000.0	0.000458	0.013	434000.0
0.0	2250.0	10000.0	0.000362	0.01	902000.0
0.0	2500.0	10000.0	0.000293	0.015	499000.0
0.0	2750.0	10000.0	0.000242	0.016	555000.0
0.0	3000.0	10000.0	0.000204	0.012	1210000.0
1.0	500.0	100.0	0.0733	0.763	8.23
1.0	750.0	100.0	0.0326	0.757	18.8
1.0	1000.0	100.0	0.0183	0.631	48.0
1.0	1250.0	100.0	0.0117	0.446	151.0
1.0	1500.0	100.0	0.00814	0.445	218.0
1.0	1750.0	100.0	0.00598	0.417	338.0
1.0	2000.0	100.0	0.00458	0.367	569.0
1.0	500.0	1000.0	0.0232	0.77	8.08
1.0	750.0	1000.0	0.0103	0.46	50.9
1.0	1000.0	1000.0	0.00579	0.412	113.0

Continued on next page

Table A.1 – continued from previous page

$\cos\iota$	f_0 [Hz]	τ [s]	ϵ	$d^{90\%}$ [Mpc]	$E^{90\%}$ [$M_\odot c^2$]
1.0	1250.0	1000.0	0.00371	0.357	235.0
1.0	1500.0	1000.0	0.00257	0.267	603.0
1.0	1750.0	1000.0	0.00189	0.165	2160.0
1.0	2000.0	1000.0	0.00145	0.188	2180.0
1.0	2250.0	1000.0	0.00114	0.039	62400.0
1.0	2500.0	1000.0	0.000927	0.042	69200.0
1.0	2750.0	1000.0	0.000766	0.051	56200.0
1.0	3000.0	1000.0	0.000644	0.048	74300.0
1.0	500.0	10000.0	0.00733	0.402	29.7
1.0	750.0	10000.0	0.00326	0.275	143.0
1.0	1000.0	10000.0	0.00183	0.216	411.0
1.0	1250.0	10000.0	0.00117	0.181	918.0
1.0	1500.0	10000.0	0.000814	0.164	1610.0
1.0	1750.0	10000.0	0.000598	0.123	3860.0
1.0	2000.0	10000.0	0.000458	0.092	9040.0
1.0	2250.0	10000.0	0.000362	0.048	41400.0
1.0	2500.0	10000.0	0.000293	0.053	43400.0
1.0	2750.0	10000.0	0.000242	0.047	65400.0
1.0	3000.0	10000.0	0.000204	0.079	27500.0

Appendix B

The noise covariance of the TF-maps

This appendix studies the covariance of the simulated noise used in Ch. 4, and validity of the using only correlations diagonal in frequency and time in the likelihood.

For the first test, $\mathcal{O}(10,000)$ s of stationary aLIGO colored Gaussian noise was simulated, and TF-maps created with 4-second Tukey-windowed FFTs as in Sec. 4.4. Histograms were then made measuring the different types of correlations that are listed below:

- Correlations between pixels of the same frequency but at neighboring time-segments - i.e., time-segments adjacent to each other in the TF-map
- Correlations between distinct frequency bins at neighboring time-segments
- Correlations between frequency bins across all time-segments (i.e., not just neighboring ones)
- Finally for completeness, pixels at different frequencies at the same time-segments

Example histograms are plotted in Fig. B.1 for these different types of correlations. We see that the mean values of the noise-correlation histograms are consistent with zero. This is expected when any correlations between these frequency bins across time

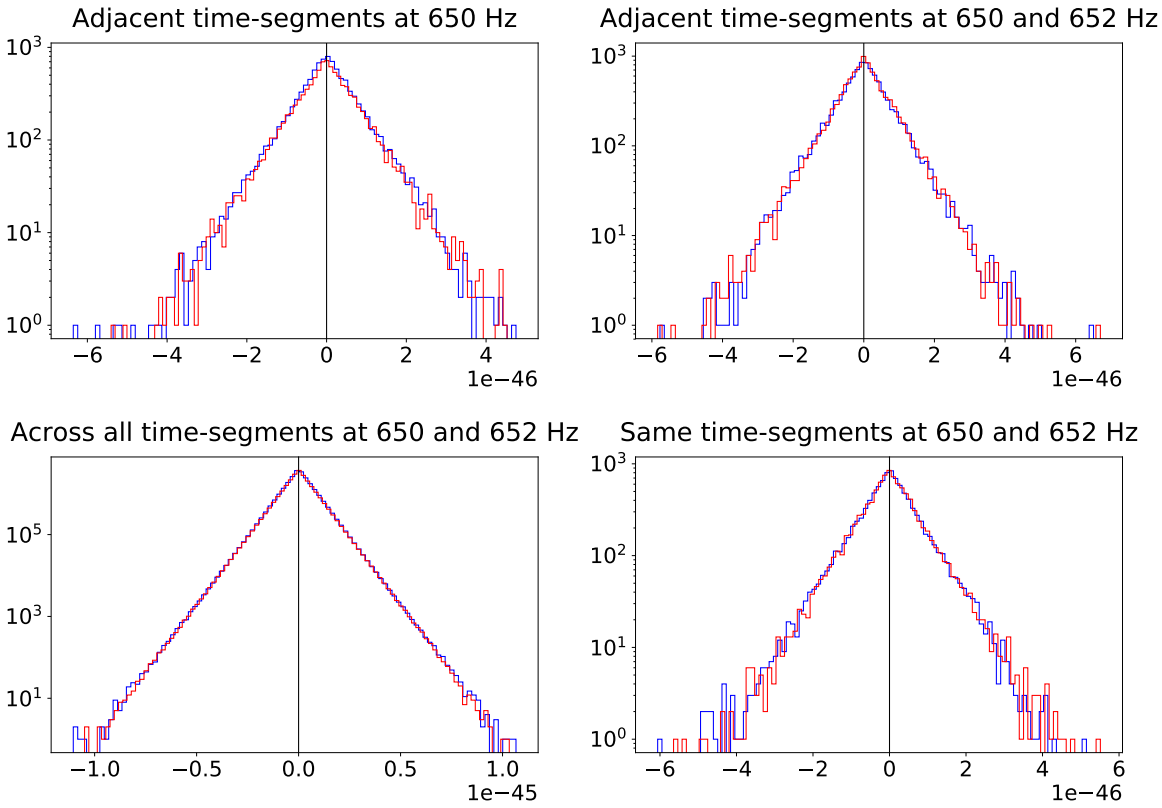


Figure B.1: Starting from the top left in clockwise direction, the plots show histograms of (i) Correlations between two bins at 650 Hz at adjacent time-segments, (ii) Correlations between bins at 650 and 652 Hz at adjacent time-segments, (iii) Correlations between bins at 650 and 652 Hz across all time-segments and (iv) Correlations between bins at 650 and 652 Hz at the same time-segments. The blue and the red traces correspond to the real and imaginary parts of the correlations. The frequencies of 650 and 652 Hz are chosen as examples. The absolute values of the means of the correlations are (clockwise from top-left) $3.3 \times 10^{-49}, 8.8 \times 10^{-49}, 6.5 \times 10^{-52}, 1.7 \times 10^{-48}$. Reproduced from Ref. [11] with permission.

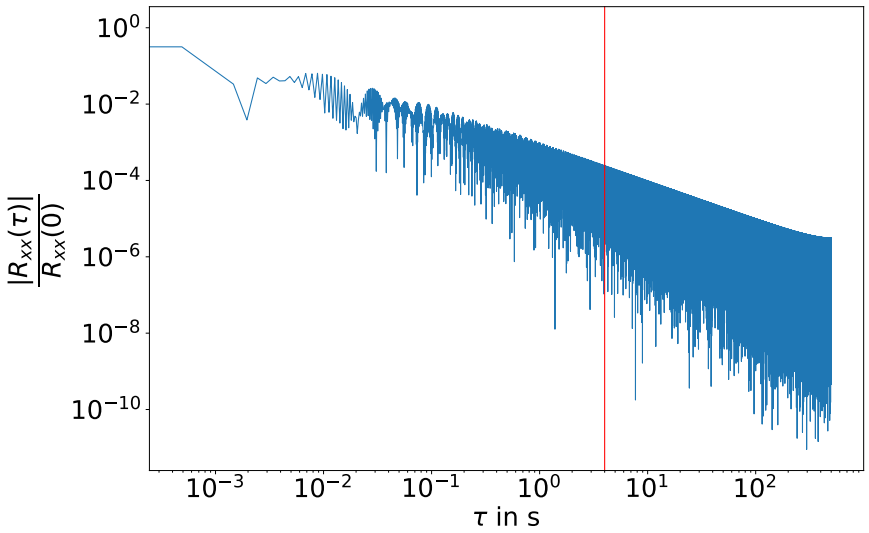


Figure B.2: The absolute value of the autocorrelation function is plotted as a function of the time difference. The correlation value at four seconds - indicated by the red line - is at least two orders of magnitude smaller than 1. Reproduced from Ref. [11] with permission.

segments are negligible. This was true across different frequency choices in the advanced LIGO sensitive frequency band from $\sim 20 - 1000$ Hz.

Another test was performed to study correlations across time using the time-domain auto-correlation function. The noise correlation duration is used as a metric to measure the duration of correlations of stationary (ergodic) random data. It is defined as [284]:

$$T_n = \frac{2}{R_{xx}(0)} \int_0^\infty |R_{xx}(\tau)|\, d\tau, \tag{B.1}$$

where $R_{xx}(\tau)$ is the time-domain auto-correlation function calculated as the inverse Fourier transform of a two-sided PSD. When using the O2 PSD of advanced LIGO without narrow-band features, and with a sampling frequency of 2048 Hz (and $f_{max} =$ 1024 Hz) we get $T_n \approx 0.34$ seconds. This is already more than an order of magnitude smaller than the segment duration of 4s used in Ch. 4, but most correlations come from the lowest frequencies that are not typically used in gravitational-wave analysis. Indeed if we only consider frequencies above 30 Hz the correlation duration drops to around 0.02 seconds.

Finally, the absolute value of the autocorrelation function with a low-frequency cut-off of 30 Hz is plotted in Fig. B.2. This shows that the correlations at $\tau = 4$ seconds are several orders of magnitude smaller in magnitude than $R_{xx}(\tau = 0)$. We should stress again that these tests are valid only for noise that is stationary and Gaussian for the duration of the entire spectrogram without strong narrow-band features, assumptions that would not generally be true with real data.

Appendix C

Coherent phase marginalization

Chapter 4 developed methods to marginalize over the complex phase of the signal to handle certain modeling uncertainties. However, the likelihood expression derived there was phase incoherent, i.e., the signal phase in each detector was marginalized over independently. A better way to treat multiple detectors is to account for the fact that the phase from a common GW signal will be coherent across detectors in a manner that depends on their sky position. This appendix derives an expression for such a coherently marginalized likelihood.

Let us suppose we have a network of N detectors. Let us start with the multiple detector likelihood similar to the one in Eq. 4.5 for a single time-frequency bin or a pixel.

$$\mathcal{L}(\tilde{d}_{ij}|\boldsymbol{\theta}, \phi^h_{ij}) = \prod_k \frac{1}{2\pi S^n_{ijk}} \exp\left(-\frac{2}{T}\frac{|\tilde{s}_{ijk}|^2 + |\tilde{h}_{ijk}|^2}{S^n_{jk}}\right) \exp\left(\frac{4}{T}\frac{|\tilde{s}_{ijk}||\tilde{h}_{jk}|\cos(\phi^s_{ijk} - \phi^h_{ijk})}{S^n_{jk}}\right), \tag{C.1}$$

where k is an index over detectors. Now let the phase of the GW at the center of the Earth (in the $i-j$ pixel be) be ϕ^h_{ij}. The phase at the k-th detector will be (see Eq. 1.12)

$$\phi^h_{ijk} = \phi^h_{ij} + \frac{2\pi f}{c}\hat{\mathbf{n}} \cdot \hat{\mathbf{x}}_k, \tag{C.2}$$

where $\hat{\mathbf{x}}_k$ is the unit vector for the position of the k^{th} detector from the center of the Earth. This equation demands that – while we may want to marginalize over the

unknown phase – it also should be consistent between detectors.

Upon repeating this for all the N detectors, we can write Eq. C.1 as,

$$\begin{aligned}\mathcal{L}(\tilde{d}_{ij}|\boldsymbol{\theta},\phi^h_{ij}) =& \exp\left[\sum_k z_{ijk}\cos(\phi^s_{ijk}-\phi^h_{ijk})\right] \\ &\times\prod_k \frac{1}{2\pi S^n_{ijk}}\exp\left(-\frac{2}{T}\frac{|\tilde{s}_{ijk}|^2+|\tilde{h}_{ijk}|^2}{S^n_{jk}}\right),\end{aligned} \tag{C.3}$$

where,

$$z_{ijk} = \frac{4|\tilde{d}_{ijk}||\tilde{h}_{ijk}|}{TS^n_{jk}}. \tag{C.4}$$

The phase-marginalization integral in this case, the equivalent of Eq. 4.6, is given by

$$I^{\phi}_{ij}(\hat{\mathbf{n}}) = \frac{1}{2\pi}\int_0^{2\pi} d\phi^h_{ij}\exp\left[\sum_k z_{ijk}\cos(\phi^s_{ijk}-\phi^h_{ijk})\right]. \tag{C.5}$$

Using C.2 and defining $\delta_{ijk} = \phi^s_{ijk} - \frac{2\pi f}{c}\hat{\mathbf{n}}\cdot\hat{\mathbf{x}}_k$, we get

$$I^{\phi}_{ij}(\hat{\mathbf{n}}) = \frac{1}{2\pi}\int_0^{2\pi} d\phi^h_{ij}\exp\left[\sum_k z_{ijk}\cos(\delta_{ijk}-\phi^h_{ij})\right]. \tag{C.6}$$

Note that δ_{ijk} is dependent only on the direction $\hat{\mathbf{n}}$ and is a constant for the purpose of this integral. With some trignometric expansion we get,

$$I^{\phi}_{ij}(\hat{\mathbf{n}}) = \frac{1}{2\pi}\int_0^{2\pi} d\phi^h_{ij}\exp\left[\cos\phi^h_{ij}\left(\sum_k z_{ijk}\cos\delta_{ijk}\right)+\sin\phi^h_{ij}\left(\sum_k z_{ijk}\sin\delta_{ijk}\right)\right]. \tag{C.7}$$

This integral reduces to a zeroth order modified Bessel function again and can be written as

$$I^{\phi}_{ij}(\hat{\mathbf{n}}) = I_0\left[\sqrt{\left(\sum_k z_{ijk}\cos\delta_{ijk}\right)^2+\left(\sum_k z_{ijk}\sin\delta_{ijk}\right)^2}\right]. \tag{C.8}$$

Finally, this gives the phase-marginalized likelihood

$$\begin{aligned}\mathcal{L}(\tilde{d}_{ij}|\boldsymbol{\theta},\phi_{ij}^{h}) =& I_0\left[\sqrt{\left(\sum_k z_{ijk}\cos\delta_{ijk}\right)^2+\left(\sum_k z_{ijk}\sin\delta_{ijk}\right)^2}\right] \\ &\times\prod_k \frac{1}{2\pi S_{ijk}^n}\exp\left(-\frac{2}{T}\frac{|\tilde{s}_{ijk}|^2+|\tilde{h}_{ijk}|^2}{S_{jk}^n}\right).\end{aligned} \tag{C.9}$$

Appendix D

Acronyms

This appendix defines some of the acronyms that have been used through this dissertation, roughly in the same order as they first occur in the text.

D.1 Acronyms

Table D.1: Acronyms

Acronym	Meaning
GW	Gravitational waves.
LIGO	Laser Interferometer Gravitational-Wave Observatory.
aLIGO	Advanced Laser Interferometer Gravitational-Wave Observatory, representing the second state of the LIGO that came online in 2015.
aVirgo	Advanced Virgo, representing the second state of the Virgo that came online in 2017.
BBH	Binary black hole.
BNS	Binary neutron star.

Continued on next page

Table D.1 – Continued from previous page

Acronym	Meaning
O1	The first observing run of the advanced detectors, from September 12th, 2015 to January 19, 2016. Only the aLIGO detectors were taking science data during this run.
O2	The first observing run of the advanced detectors, from Nov 30th, 2016 to August 25th, 2017. The aVirgo detector joined the run on August 1, 2017
O3	The third observing run of the advanced detectors, from April 1st, 2019 to March 27th, 2020.
O3a	The first half of O3 from April 1st, 2019 to October 1st 2019.
O3b	The second half of O3 from November 1st, 2019 to March 27th, 2020.
LISA	Laser Interferometer Space Antenna.
TT	Transverse traceless.
CBC	Compact body coalescences.
DWD	Double white dwarf.
EMRI	Extreme mass ratio inspiral.
SGWB	Stochastic gravitational-wave background.
KAGRA	Kamioka Gravitational-Wave Detector.
SQL	Standard quantum limit.
DARM	Differential arm.
INTEGRAL	INTErnational Gamma-Ray Astrophysics Laboratory.
PSD	Power spectral density.
MCMC	Markov Chain Monte Carlo.
MH	Metropolis-Hastings (algorithm).
STAMP	Stochastic Transient Analysis Multidetector Pipeline.
TOV	Tolman-Oppenheimer- Volkoff (limit).
FT-map	Frequency time maps.
SNR	Signal-to-noise ratio.

Continued on next page

Table D.1 – Continued from previous page

Acronym	Meaning
FFT	Fast Fourier transform.
H1	The LIGO Hanford detector.
L1	The LIGO Livingston detector.
HMM	Hidden Markov model.
GWTC-1	Gravitational-wave transient catalog 1, containing list of GW detections from O1 and O2.
GWTC-2	Gravitational-wave transient catalog 1, containing list of GW detections from O1 and O3a.
EM	Electromagnetic.
SDSS	Sloan digital sky survey .
SGCN	Stochastic gravitational-wave confusion noise. A common term to refer to both SGWBs and stochastic gravitational wave foregrounds.
PTA	Pular timing array.
BLIP	Bayesian LISA Pipeline.
TDI	Time delay interferometry.